

Rhode Island College Digital Commons @ RIC

Faculty Publications

11-2000

Staff Development Strategies and Contexts Associated with Positive Impacts on Teachers'— Attitudes and Practices

Susan Gracia

Rhode Island College, sgracia@ric.edu

Follow this and additional works at: <https://digitalcommons.ric.edu/facultypublications>



Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Teacher Education and Professional Development Commons](#)

Citation

Gracia, Susan, "Staff Development Strategies and Contexts Associated with Positive Impacts on Teachers' Attitudes and Practices" (2000). *Faculty Publications*. 270.

<https://digitalcommons.ric.edu/facultypublications/270>

This Dissertation is brought to you for free and open access by Digital Commons @ RIC. It has been accepted for inclusion in Faculty Publications by an authorized administrator of Digital Commons @ RIC. For more information, please contact digitalcommons@ric.edu.

Boston College

Lynch School of Education

Department of Educational Research,
Measurement, and Evaluation

Educational Research, Measurement,
and Evaluation Program

STAFF DEVELOPMENT STRATEGIES AND CONTEXTS
ASSOCIATED WITH POSITIVE IMPACTS ON
TEACHERS' ATTITUDES AND PRACTICES

Dissertation

by

SUSAN M. GRACIA

submitted in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

November 2000

Abstract

Staff Development Strategies, Organizational Contexts, and Teacher Characteristics Associated with Positive Impacts on Teachers' Attitudes and Practices

by

Susan Gracia

Dissertation Advisor: Dr. Joseph Pedulla

The purpose of this study was to identify staff development strategies, organizational contexts, and teacher characteristics that were associated with positive impacts on teachers' attitudes and practices. The study focused on 89 elementary teachers in 11 schools across 5 districts, all of whom participated in a state-sponsored staff development project during the 1999/2000 academic year. This staff development initiative was designed to assist teachers in acquiring the knowledge and skills necessary to develop and implement standards-based math instruction and assessment in their classrooms.

Using surveys, interviews, and written teacher logs, data were gathered regarding the amounts and types of staff development in which teachers engaged, teachers' concerns about standards-based instruction and assessment, school and teacher characteristics thought to influence the adoption and use of standards-based instruction and assessment, and the frequency with which teachers employed various instructional and assessment strategies in their classrooms. Stepwise regression analyses were subsequently performed to explore the utility of using school features, teacher

characteristics, and types of staff development in which teachers engaged to predict residual gains in concern about and reported use of standards-based instruction and assessment.

Study findings revealed that residual gains in level of concern about standards-based instruction and assessment in mathematics were significantly predicted by teacher involvement in the following staff development activities: examining/analyzing student work and implementing standards-based classroom activities while carrying out action research. Furthermore, increases in teacher self-efficacy were found to be associated with residual gains in intensity of concern, while involvement in curriculum development was associated with negative gains in level of concern. Finally, analyses revealed that a combination of personal (years teaching), organizational (principal supportiveness and organizational climate), and staff development (curriculum development) factors were influential in predicting residual gains in use of standards-based instruction and assessment.

The implications of this study on the design of future change initiatives were presented, as were recommendations for future research. The limitations of this study were also thoroughly discussed.

Acknowledgements

Several people have helped me in many ways to complete this dissertation. In particular, I would like to thank Dr. Joseph Pedulla, chairman of the Dissertation Committee, who helped me design this study and set a realistic timeframe for completing it. He also set high standards for the quality of my work, gave me useful advice, and was available for questions whenever I had them. As a result, I am very grateful for his guidance and assistance. The other members of my dissertation committee, Dr. Michael Russell and Dr. Irwin Blumer, were also instrumental to the success of this project. Dr. Russell's knowledge of the staff development project on which this study focuses, as well as his technical expertise, enabled him to offer me very detailed, useful feedback that I in turn utilized to improve the study. Dr. Blumer's thorough reading of the various drafts of my dissertation, as well as his probing questions and suggestions for additional research I should consult, also helped improve and shape the final product.

I am also very grateful for the constant support of my husband, Alvaro Gracia. He wanted me to finish this project as much as I did, not because he wanted it over with, but because he felt that it would be such a wonderful accomplishment for me. He was also very generous in helping me secure the time and resources necessary to complete this dissertation. In addition, I will always be grateful to my children, Christian and Claire, who are too young right now to understand what I have done but were a constant source of inspiration for me. I am also appreciative of the confidence my mother, Susan Metzger, has always had in me, as well as her willingness to proofread the final draft of my dissertation.

Finally, I would also like to recognize the influence of my late father, James Metzger, who was very supportive of me in the first two years of my Ph.D. program and would have been so happy and proud to see me complete it.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	I
LIST OF TABLES.....	VI
LIST OF FIGURES	VIII
CHAPTER 1	1
Statement of problem.....	1
Background.....	3
Staff development as a tool of reform.....	4
Research on effective staff development strategies	6
Challenging the dominant training model.....	8
A new staff development paradigm.....	13
Research Questions.....	16
Summary and Conclusions	17
CHAPTER 2	18
Review of the Literature.....	18
Overview	18
Theories of Change.....	18
Generalized change theories	18
Educational change theories and approaches	21
Instrumental-rational educational change models	21
Normative-cultural educational change models	23
Personal-transformative educational change models.....	29
A Brief History of Staff Development for Teachers in the United States.....	37
Dominant 20 th Century Staff Development Models.....	46
Staff development “models” as perspectives from which to view and understand staff development .	46
Staff development “models” as specific staff development strategies and techniques	48
Training Model	49
Research on the Training Model	50
Individually Guided Staff Development Model	56
Research On Individually Guided Staff Development.....	57
Observation/assessment model	60
Research on the Observation/Assessment Model.....	61
Action research model	63
Research on the Action Research Model	64
Organizational development (OD) model.....	66

Research on the Organizational Development Model	67
Development/improvement process model	69
Research on the Development/Improvement Process Model	72
Professional Development School model	77
Research on the Professional Development School model	78
Organizational context model	80
Research on the Organizational Context Model	81
Gaps in the Staff Development Research	86
Summary	91
 CHAPTER 3	 92
Research Design.....	92
Overview	92
Research Hypotheses	92
Independent Variables.....	95
Dependent Variables	99
Population	102
Teachers	102
Students	103
Subjects	104
Teachers	104
Students	105
Methods.....	105
Teacher survey: pre	105
Teacher survey: post.....	109
Professional development activity logs	109
Student surveys	110
Teacher Interviews.....	110
Collection of Data.....	111
Teacher survey: pre	111
Professional development activity logs	112
Teacher interviews	112
Student surveys	113
Treatment of Data	114
Teacher survey data: pre.....	114
Teacher survey: post	115
Professional development activity logs	116
Interview data	117
Student survey data	117
Statistical methods	117
Limitations	120
Summary	122
 CHAPTER 4	 123
Results.....	123
Response Rates and Sample Size.....	123
Descriptive Statistics.....	124
Independent Variables	124

Dependent Variables	131
Student surveys	136
Variables Excluded from Analysis	137
Analysis of the Relationships among Independent Variables and Residual Gains in Teachers' Level of Concern	139
Analysis of the Relationships among the Independent Variables and Residual Gains an Teachers' Intensity of Concern	154
Analysis of the Relationships among the Independent Variables and Residual Gains in Teachers' Reported Use	164
Analysis of Teacher Interviews	176
Implementation of Standards-Based Activities in the Classroom	178
Work with Educational Consultants	186
Curriculum Development	190
Workshops	193
Action Research	195
Meetings/Discussions with Colleagues	199
Examining Student Work	201
Communication and Organizational Climate	203
Quality of School/Community Relations	209
Principal Directiveness and Supportiveness	212
Years Teaching	214
Summary of Results	216
CHAPTER 5	219
Conclusion and Discussion	219
Summary	219
Residual Gain in Teachers' Level of Concern	219
Residual Gain in Teachers' Intensity of Concern	222
Residual Gain in Teachers' Frequency of Use of Standards-Based Instruction and Assessment	223
Contributions of the Research	225
Implications of Study Findings	227
Limitations	232
Suggestions for Future Research	234
Conclusions	235
BIBLIOGRAPHY	238
APPENDIX 1: TEACHER PRE- AND POST-SURVEYS	251
APPENDIX 2: PROFESSIONAL DEVELOPMENT ACTIVITY LOG	252
APPENDIX 3: STUDENT SURVEY	253
APPENDIX 4: INTERVIEW PROTOCOL	254

LIST OF TABLES

Table		Page
1	Number of Survey Items Per Variable.....	106
2	Coefficients of Internal Reliability for the SoC Questionnaire, N=830.....	107
3	School Sizes	125
4	Descriptive Statistics: General Self-Efficacy, School-Wide Communication, Principal Directiveness, Principal Supportiveness, School/Community Relations, and Organizational Climate (N=89)	126
5	Staff Development Activities Engaged in by Study Participants.....	130
6	Level of Concern Results from Pre- and Post-SoC Surveys.....	132
7	Pre, Post, and Residual Gains in Level of Concern.....	133
8	Pre, Post, and Residual Gains in Intensity of Concern.....	134
9	Pre, Post, and Residual Gains in Use of Standards-Based Instruction and Assessment.....	136
10	Reliability Estimates of Self-Efficacy, Principal Directiveness, Principal Supportiveness, School/ Community Relations, Organizational Climate, and School-Wide Communication Scales.....	138
11	Zero Order Correlations among Criterion and Predictor Variables.....	140
12	Casewise Diagnostics (Dependent Variable= Residual Gain in Level of Concern)	142
13	Summary of Stepwise Regression Analysis for Variables Predicting Residual Gains in Level of Concern (N=89)	146

LIST OF TABLES

Table		Page
14	Summary of Stepwise Regression Analysis for Independent Variables and Interaction Variables Predicting Residual Gains in Level of Concern.....	148
15	Regression Coefficients, Standard Errors, and T-Values for High, Average, and Low Rates of Implementation of Standards-Based Activities in the Classroom.....	151
16	Casewise Diagnostics (Dependent Variable= Residual Gain in Intensity of Concern)	157
17	Summary of Stepwise Regression Analysis for Variables Predicting Residual Gains in Intensity of Concern (N=89)	160
18	Casewise Diagnostics (Dependent Variable= Residual Gain in Reported Use of Standards-Based Instruction and Assessment)	166
19	Summary of Stepwise Regression Analysis for Variables Predicting Residual Gains in Use of Standards-Based Instruction and Assessment (N=70).....	170
20	Regression Coefficients, Standard Errors, T-Values, and Significant Test Results for the Interaction Model (N=70)	172
21	Number of Times Teachers Were Interviewed.....	177

LIST OF FIGURES

Figure		Page
1	Normal probability plot of standardized residuals when independent variables are regressed on gain in level of concern.....	143
2	Plot of studentized residuals against standardized predicted values of gain in level of concern.....	144
3	Disordinal interaction between the action research and standards-based classroom activities variables.....	154
4	Normal probability plot of standardized residuals when independent variables are regressed on gain in intensity of concern	158
5	Plot of studentized residuals against standardized predicted values of gain in intensity of concern	158
6	Normal probability plot of standardized residuals when independent variables are regressed on gain in reported use	167
7	Plot of studentized residuals against standardized predicted values of gain in reported use	167

CHAPTER 1

Statement of problem

Staff development for teachers -- also known as continuing education, in-service training, and professional development -- is viewed by policy makers, school administrators, and the public as a vital tool for instituting school change. A component of virtually every educational innovation, staff development is generally designed to “alter the professional practices, beliefs, and understanding” of teachers in an effort to bring about an improvement in student learning (Guskey, 1986, p. 5).

Despite the prevalence and widespread acceptance of staff development, there is very little empirical evidence regarding the types of staff development strategies which are most effective at altering teachers’ beliefs and practices. Very few staff development efforts are even evaluated (Gordon, 1974). Furthermore, when staff development initiatives are evaluated, program effectiveness tends to be judged in terms of participant satisfaction, changes in knowledge, or process measures. Only rarely are changes in participants’ attitudes and practices taken into account.

Until fairly recently, the dominant staff development model utilized in U.S. educational circles was the training model. As a result, the majority of staff development research has focused on the effects of training on teachers’ attitudes, knowledge, and, albeit infrequently, practices. Many researchers who have attempted to identify the critical elements of successful staff development programs have relied on the use of meta-analysis of evaluation reports and research studies, with overly broad and often

contradictory conclusions (Guskey & Huberman, 1995). Furthermore, the results of meta-analyses very consistently contradict the findings of primary research studies.

Many staff developers and researchers have suggested alternative staff development models, such as observation/assessment, action research, individually-guided staff development, organizational development approaches, the development/improvement process, and the organizational context model. Unfortunately, few large-scale studies have been conducted on the effects of staff development models other than the training paradigm, and the amount and types of staff development activities necessary to accomplish a given purpose are virtually unknown (Stout, 1996). In addition, research on alternative staff development models has produced rich data about implementation and change efforts in single settings. However, this research is not, on the whole, generalizable to other contexts (Guskey, 1997).

This state of affairs has led some staff development experts to conclude that there is no “one size fits all” approach to staff development and that staff development must be configured differently every time it is implemented, depending on the context and personalities of those involved. However, the scarcity of evidence pointing to the types and amounts of staff development necessary to produce change makes it very difficult for staff developers to design programs that are appropriate for different contexts.

A basic, unanswered question in the field of staff development, therefore, is: Which professional development strategies and contexts have a positive impact on teachers’ attitudes and practices? Clearly, further research is needed to furnish scientific evidence of the effects of particular staff development strategies on teachers’ attitudes

and classroom practices *in a variety of contexts*. In this way, substantive conclusions concerning the types and intensity of various staff development practices may be reached. In addition, this research will allow change agents to predict (and hence design) the ideal staff development program for a particular context.

Background

The last two decades have been characterized by a near constant focus on school reform. The roots of school reform in the 1980s and 1990s lie not only in a generalized concern with schools' deficiencies, but also in concerns related to American competitiveness in a global economy, the need to address demographic shifts, an "unraveling of the social fabric," cynicism about the government, and changing political philosophies (Lashway, 1999; Fenstermacher & Berliner, 1985).

According to Lashway (1999), three waves of reform have been witnessed since 1980. The first wave, known by some as the "intensification era" (1980-1987), focused on top-down government efforts to tighten control of education. During the "restructuring era" (1988-1995), emphasis was placed on decentralization, professional empowerment, and consumer choice. In contrast, the focal points of the present "restructuring era" (1996-present) are standards, accountability, and privatization.

Furthermore, five "streams" of reform characterize the present era:

- *Reforms in subject matter standards, curriculum, and pedagogy* aspire toward more ambitious student outcomes and require a fundamental change in the nature of students' intellectual tasks and teacher-student relations

- *Reforms centered on problems of equity among a diverse student population* require teachers to identify and alter classroom practices that contribute to student failure and that undermine equal opportunity to learn
- *Reforms in the nature, extent, and uses of student assessment* require widespread, rigorous use of authentic assessment on the part of teachers
- *Reforms in the social organization of schooling* necessitate broad school restructuring oriented toward new principles
- *Reforms in the professionalization of teaching* center on teachers' demonstrated knowledge base, on conditions surrounding teacher certification and licensure, and on the structure of career opportunities in teaching (Little, 1993, pp. 130-132)

Staff development as a tool of reform

As evidenced in the five points above, teachers are frequently placed at the center of reform. For good reasons and bad, they are often viewed both as the source of the ills of today's schools and as the solution to most educational problems. However, implementation of reform presents new challenges to teachers:

To be certain, reforms pose certain technical demands—demands on the knowledge, skill, judgment, and imagination of individuals. In that sense, the implementation problem at the level of the classroom is real (Little, 1993, p. 129).

Consequently, staff development has come to play a vital role in the implementation of school reform efforts.

Staff development for teachers -- also known as continuing education, in-service training, and professional development -- is viewed by policy makers, school administrators, and the public as a vital tool for instituting school reform. A component

of virtually every educational innovation, staff development is generally designed to “alter the professional practices, beliefs, and understanding” of teachers in an effort to bring about an improvement in student learning (Guskey, 1986, p. 5). Substantial expenditures of federal and local funds support staff development (e.g., \$386 million alone in California in 1988), the majority of which are invested toward school improvement efforts.

In the United States, staff development for teachers began with Teacher Institutes in the early 19th century (Guskey, 1986). Federal interest in staff development peaked between 1956 and 1975, when federal intervention in training was used as a means of improving curricula and teaching skills. Today, state and local policy makers are responding to calls for school reform by instituting school change initiatives with staff development at their core. As a result, staff development initiatives continue to expand based on a common assumption of benefit to the public (Stout, 1996).

According to Stout (1996), the market system for providing staff development is similar to a “giant academic bazaar.” Thousands of potential programs are available to educators and administrators searching for a staff development program. Colleges, universities, local education agencies, county and state education agencies, private consultants, and publishers and manufacturers of instructional materials all compete in the staff development marketplace. With so many staff development providers, any single training topic may potentially be addressed using a multitude of different methods, materials, and timeframes, in differing settings, and at varying costs. With so many options, it can be difficult for school personnel to determine which staff development

program is most effective at helping teachers develop new skills, attitudes, and knowledge or which approach is most suitable for a particular context or purpose.

Research on effective staff development strategies

An examination of the literature on staff development is not always helpful in identifying the best staff development model for a particular purpose. One explanation for this is the fact that most publications concerning staff development are evaluation reports rather than research. Control groups are rarely used, and evaluation reports tend to be produced only for “successful” programs. Furthermore, very few school systems or individual institutions even evaluate their staff development efforts (Gordon, 1974). For example, a review of staff development patterns and policies in California concluded that staff development is rarely evaluated for the importance or coherence of its overall program goals, for the relationship between staff development goals and other school improvement goals, for the match between goals and strategies, or for the impact of staff development at the classroom level (Little, 1987). As a result, little is known about the effectiveness of the majority of staff development initiatives.

When evaluation of staff development does take place, program effectiveness is most often judged by an index of participants’ satisfaction with the program or some indication of a change in professional knowledge (Guskey, 1995). Alternatively, staff development may be evaluated by participation rates or other process measures (Little et al., 1987). Rarely do evaluations focus on teachers’ actual use of the knowledge and skills they have gained or the impact of these changes on student achievement. In addition, measurement techniques tend to be “fuzzy,” consisting of subjective opinions

such as, “Teachers felt that the program helped them improve their classroom questioning techniques” (Sparks, 1983), or based on tests that staff developers themselves created.

Given the lack of convincing evaluation data concerning staff development, the school administrator seeking information on staff development practices which have been proven effective may choose instead to consult “real” research on the subject. According to some researchers (McLaughlin & Berman, 1977; Sparks, 1987; Swenson, 1982), however, staff development as a field has few theoretical or conceptual roots and a generally meager research base. Very few “true” experiments have been conducted on the process of staff development training and its impact on everyday teaching practice. On the contrary, existing research is heavily based on descriptive surveys of the desired content and procedures for the delivery of staff development, as described by teachers and administrators (Daresh, 1985), rather than the impact of staff development strategies on teacher attitude and practice.

In 1957, approximately 50 studies (including 6 experimental studies) had been conducted in the area of training, curriculum improvement, or the implementation of innovations. By 1977, the knowledge base had broadened considerably, but nearly all literature was descriptive or conceptual. In fact, only a small proportion of articles or books even reported original research or mentioned existing studies as of 1977 (Showers et al., 1987).

Between 1977 and 1984, more than 400 studies of staff development were conducted. According to Daresh (1985), 60% of these studies were descriptive studies or survey research, while 25% were quasi-experimental. As is typical of research in the

social sciences, none of these studies utilized a true experimental design. Questionnaires were the sole data collection method in 75% of these studies. More than one data collection technique was used in fewer than 1 study out of 4. One hundred sixty studies focused on participants' preferred training topics or content, while 180 studies addressed the preferred methods of planning and carrying out staff development. Only 60 studies investigated the effects of staff development on the experiences of teachers or other school staff. What's more, results of these 60 studies indicated that staff development, as it was generally carried out in that era, had "little or no discernable effect on the attitudes or observable behaviors, at least on a short-term basis," of teachers (Daresh, 1985, p.8)! No studies looked at the long-term effects of training.

Challenging the dominant training model

It should be pointed out that until the 1980s, the dominant staff development model was the training model. This model, which focused primarily on expanding teachers' individual repertoires of skillful classroom practice, tended to offer standardized solutions (in the form of workshop series, special courses, or in-service days devoted to transmitting specific ideas, practices, or materials to teachers) to the demands of school reform and other school-level change efforts (Little, 1993). According to Wood & Thompson (1993), staff development was designed and implemented based on the following, untested assumptions well into the 1980s:

- no changes in teaching practice require more than a year to implement
- reform does not take much time; therefore, 2-3 in-service days/year are sufficient

- inspirational speakers at one-shot sessions are the best way to get teachers excited about new ideas, programs, procedures
- teachers will automatically transfer what they learn in workshops into the classroom without assistance
- teachers should sit and listen during training; movement, discussion, or practice will disturb learning
- teachers do not want to be responsible for their own professional growth
- leadership for in-service activities should come from outside consultants or administrators, and not from teachers themselves
- the primary focus of staff development should be district wide because schools within a district are identical and have the same problems

A few, noteworthy studies at this time served to challenge these assumptions and explain why the 60 studies of the impact of staff development on teacher practice produced such dismal results (i.e., that prevailing staff development strategies were not effective). One study that is cited again and again in the staff development literature is the 4-year, 2-phase Rand study (completed in 1978) which examined approximately 300 educational innovations to determine why some innovative projects succeed and others fail. In all, 852 administrators and 689 teachers were surveyed, and researchers observed programs in operation at two different time points. The Rand study (Berman & McLaughlin, 1978) concluded that certain staff development strategies had great impact on the success of innovations:

- training that is concrete, on-going, and teacher-specific produces changes in classroom practices; one-shot workshops are not effective
- local resource personnel who provide "on-call" advice are more effective than outside consultants whose advice is too general, untimely, and irrelevant
- opportunities to observe projects in other classrooms or districts are useful to teachers because they receive advice and encouragement from peers who have experienced success in implementing innovations
- staff support activities are important-e.g., regular project meetings to discuss and work on problems
- principal participation is vital to help teachers with program objectives and to support their efforts

Clearly, the Rand study was one of the first indications that the prevailing theory of staff development was flawed.

Rather than conduct original research of their own, many researchers in the 1970s and 80s attempted to identify the critical elements of successful staff development programs through the use of meta-analysis. These studies were conducted with mixed results. On the one hand, meta-analyses focused solely on main effects, or components or processes that were consistent across staff development programs and contexts (Guskey, 1997). In their search for "one right answer," many of these studies produced recommendations for successful staff development practices that were broad, causing some staff developers to complain that they were too general to be of much use to schools in designing programs.

Additionally, the contradictory conclusions reached by a number of meta-analyses have been frustrating to practitioners.. For example, McLaughlin (1990), Weatherley & Lipskey (1977), and Wise (1991) concluded that staff development efforts must be teacher-specific and focus on day-to-day activities of teachers, while others report that individual professional development efforts are detrimental (Tye & Tye, 1984; Waugh & Punch, 1987). Similarly, some meta-analyses concluded that staff development needs to be carried out by individual teachers (Joyce, McNair, Diaz, & McKibbin, 1976; Lambert, 1988; Lawrence, 1974; Massarella, 1980), in contrast to other research results suggesting that staff development be implemented under the guidance and vision of change agents who see beyond the walls of the individual classroom (Barth, 1991; Clune, 1991; Mann, 1986; Wade, 1984).

Many meta-analyses were useful, however, for shifting attention away from the dominant training paradigm and toward alternative staff development models such as observation/assessment, action research, individually-guided staff development, organizational development approaches, the development/improvement process, professional development schools, and the organizational context model. Each of these alternative staff development models is briefly defined below:

- In the *observation/assessment model*, coaches (fellow teachers or “expert” observers) visit teachers’ classrooms, observe classroom practice, and offer feedback that can be used by teachers to improve teaching and student learning.

- In the *action research model*, teachers identify a research question, collect data in their own classrooms, and make appropriate changes in classroom practice based on their analysis of the data they have gathered.
- In *individually-guided staff development*, teachers determine their own professional development goals, design their own staff development plan, and select the activities in which they participate.
- In the *organizational development approach*, school staff and facilitators work together to bring about organizational change, improve the school's organizational structure, and build internal capacity to solve future problems.
- In the *development/improvement process*, a complex model characterized by teacher participation in long term school improvement or curriculum development initiatives, any one or all of the staff development approaches discussed thus far may be utilized.
- In *professional development schools*, schools and universities collaborate to provide opportunities for novice teachers to learn from expert practitioners, while experienced teachers act as mentors, university adjuncts, school restructurers, and teacher leaders (Darling-Hammond, 1995).
- The *organizational context model* stipulates that organizational context features and teacher characteristics, rather than specific in-service activities, are responsible for the success of staff development initiatives.

This shift toward alternative staff development approaches was supported by Gordon's 1974 review of 97 staff development studies and evaluation reports. This

review suggested that the programs with the highest probability of being effective were those in which teachers planned and managed their own professional development activities, shared their experiences trying out new practices in their classrooms, and received feedback on their practices from local experts. Showers, Joyce & Bennett (1987) went a step further in their analysis of 200 research reports, concluding that, for a complex model of teaching, approximately 25 teaching episodes during which a new strategy is used are necessary for all of the conditions of transfer to be achieved. This study also found large effect sizes on knowledge acquisition as associated with expert or peer coaching; combined presentations, demonstrations, practice, and feedback; and combined theory, demonstration, practice, feedback (Showers et al., 1987).

Recently, staff development experts such as Guskey (1994) and Little (1983) have eschewed the “one size fits all” approach to the staff development research, asserting that research has neglected the uniqueness of the individual setting and teachers with different individual histories, practices, and circumstances. This is even proposed as an explanation for the opposing conclusions reached in certain meta-analyses. Guskey (1995, p. 117) summarized this view:

The combination of complexity and diversity makes it difficult, if not impossible, for researchers to come up with universal truths... We know with certainty that reforms in education today succeed to the degree that they adapt to and capitalize on this variability... Our search must focus, therefore, on finding the *optimal mix*—the assortment of professional development processes and technologies that work best in a particular setting.

A new staff development paradigm

Today, most staff development experts disavow the traditional training model.

Since the mid-1980s, there has been increasing recognition that reform of teacher knowledge, attitude, and practices calls not for training, but for adequate opportunity to learn, investigate, experiment, consult, and evaluate (Little, 1993). It has also become quite widely accepted that teachers' everyday-work with children should be at the center of staff development activities (Novick, 1996). According to Wood and Thompson (1993), staff developers and researchers (have thus developed a new set of assumptions about the staff development practices that are most likely to have an impact on teachers and teaching:

- the school should be the primary focus of improved practice and staff development
- significant changes in practice take considerable time and are the result of staff development conducted over several years
- a school culture supportive of improved practice and professional growth is essential to successful staff development
- all educators should be involved in staff development initiatives
- the principal is key to any staff development effort to improve professional practice; s/he should work with faculties to establish goals, plan staff development, participate in learning with teachers, provide follow-up assistance, and serve as advocate for teachers who are implementing more effective practices in their classrooms
- the selection of goals that guide staff development should include teachers

- change in practice is difficult and requires systematic support (coaching, feedback, etc.) to implement and sustain over time
- staff development should support instructional and program improvement and be closely linked to instructional supervision, teacher evaluation, and curriculum implementation

The redesigning of staff development and the adoption of new assumptions in the field is a sign of progress. Past research has, for the most part, demonstrated what does *not* work and has offered alternatives to the traditional training paradigm. In addition, staff development experts have focused new attention on the context in which staff development takes place and on the fact that one model may not serve all training goals, audiences, or circumstances.

However, the majority of recent research (1985-1999) conducted to identify staff development practices that are effective at changing teacher attitudes, knowledge, and practices consists of qualitative studies (Borko, Mayfield, Marion, Flexer, & Cumbo, 1997; Borko, 1997; Borko & Putnam, 1997; Flexer, Cumbo, Borko, Mayfield, & Marion, 1995; Saxe, Franke, Gearhart, Howard, & Crockett, 1997; Shepard, 1995). While these studies do present encouraging ideas regarding the implementation of effective staff development, few researchers who advocate “alternative” staff development strategies such as observation/assessment, individually-guided staff development, action research, organizational development, the development/improvement process, and the organizational context model, have attempted to show a direct relationship between these

strategies and changes in teacher's beliefs and practices. Though case studies reveal rich data about implementation and change efforts in specific contexts, they are usually conducted in a single setting and are not generalizable (Guskey, 1997), even to similar contexts. In addition, there is little literature that points to the recommended *intensity* of staff development practices that is necessary to produce an intended change.

Research Questions

Clearly, further research is needed to furnish scientific evidence of the effects of particular staff development strategies on teachers' attitudes and classroom practices *in a variety of contexts*. In this way, substantive conclusions concerning the types and intensity of various staff development practices may be reached. In addition, this research will allow change agents to predict (and hence design) the ideal staff development program for a particular context.

This study will therefore address the following research questions:

- Which staff development strategies and contexts are related to changes in teachers' concerns?
- Which staff development strategies and contexts are related to changes in teachers' classroom practices?

In order to address these questions, the present study will closely examine a large-scale staff development project, focusing on the staff development strategies employed, the contexts in which staff development takes place, and the degree of impact on teachers' attitudes and practices.

Summary and Conclusions

Staff development is critical to the implementation of many facets of school reform. Despite this fact, past research has revealed that little conclusive evidence exists regarding the types and amount of staff development that are likely to have an impact in any particular setting.

Staff development cannot continue to be implemented on a “hit or miss” basis. Staff developers owe this not only to those who invest millions of dollars in staff development, but also to the students who are expected to reap the ultimate benefits of teacher change. As with any other field of education, staff development must be held accountable for the degree to which it meets its goals. However, this will be difficult to achieve until more is known about the strategies and contexts that have a positive impact on teacher change. It is hoped that the present study will fill this gap by attempting to identify the “optimal mix” of staff development practices that is likely to impact the change process in any particular setting.

CHAPTER 2

Review of the Literature

Overview

This chapter discusses theory and research that form the background for the present investigation. This review of relevant literature is organized into sections that deal with: (a) theories of change; (b) a brief history of staff development for teachers in the United States; (c) dominant 20th century staff development models and related research; and (d) gaps in the staff development research.

Theories of Change

Generalized change theories

As discussed in the previous chapter, staff development for teachers can be viewed as strategies and activities designed to bring about *change* in the professional practices, beliefs, and understandings of educators. In order to implement staff development initiatives that produce long-lasting teacher change, therefore, school administrators and staff developers must understand the nature of change itself and how individuals are believed to respond to innovations.

The most widely accepted change theories have described change as a linear process characterized by passage through a number of “stages” or “steps.” In 1958, Lewin, for example, outlined one of the most recognized theories of change in his renowned work, Group Decision and Social Change. According to Lewin, change is effected in three stages: (1) unfreezing: breaking existing habits or customs which are the basis of an inner resistance to change; (2) moving: changing from an old situation to a

new one; and (3) refreezing: putting in place new social habits to conform with requirements of an innovation.

Lewin's model was the building block of several change models that were soon to follow. Lippit, Watson, and Westley (1958) built on Lewin's model but added several stages that would necessarily precede Lewin's "unfreezing" stage. These new change stages included: development of a need for change; establishment of a change relationship between a client and change agent; clarification or diagnosis of the client system's problem; examination of alternative routes and goals; and establishing goals and intentions of action.

Lippit, et al. (1958) identified two additional steps in the change process: transformation of intention into actual change (all of which overlap with Lewin's "moving" stage) and generalization and stabilization of change (which corresponds to Lewin's "refreezing" stage). This model improved on Lewin's model because it took into account the fact that successful change requires recognition on the part of those involved that change is needed. In 1973, Havelock devised a change model that is reminiscent of Lippit et al.'s, with one notable difference. This model stresses the fact that proposed changes must gain acceptance by those affected before being implemented.

In a similar vein, the well-known Prochaska model of behavior change repeats the notion that behavior change is more than a one-step process from "no behavior to behavior." According to the model, behavior change moves through five stages: precontemplation (no behavior), contemplation (no behavior), preparation (no behavior)

or small steps in the direction of adopting new behavior), action (behavior), and maintenance (behavior) (Prochaska, DiClemente, & Norcross, 1992).

In their four-step change model, Rogers and Shoemaker (1971) were the first to theorize a confirmation stage in which individuals seek reinforcement for the changes they have made but may reverse their changes if exposed to conflicting messages about the innovation. Zaltman, Duncan, and Holbeck (1973) subsequently distinguished between two stages in any innovation: initiation, in which participants become aware of, form an attitude, and make a decision about a proposed change, and implementation, in which an innovation is piloted and then sustained.

More recently, many of those interested in how change occurs have rejected the idea that change occurs in specific steps or stages and that careful planning of change will help ensure a smooth, predictable change process. Rather, some change theorists have embraced the notions of Chaos Theory, a theory of mathematical modeling, which assumes that “the natural condition of the physical world contains universal behaviors of complexity, randomness, jagged edges, and sudden leaps” (Wheatley, 1992, in Snyder, Acker-Hocevar, & Wolf, p. 6). According to Chaos Theory, irregularity should be embraced as the norm; hence, it is time to set aside assumptions about regularity and controllability in changing organizations and people. Change agents, therefore, should search for structures of order that exist within chaotic, unpredictable systems and guide patterns of regularity toward a vision, mission and set of goals, while at the same time viewing chaotic events as sources of energy and “exhilaration” (Snyder, Acker-Hocevar, & Wolf, 1995).

Educational change theories and approaches

Three approaches to school change have dominated US education in this century. These three views have been labeled: *instrumental-rational*, *normative-cultural*, and *personal-transformative*. In addition, each approach is characterized by one or more models of school change.

Instrumental-rational educational change models

In the instrumental-rational approach to change, the teacher is seen as a passive receiver of a change product or a reactor to change. Change is imposed from above, in response to an individual or group's conviction that there is a "right" thing to do (Rieley, 1997). As a result, participation is required, not simply voluntary, similar to the following philosophy:

The Federal Aviation Agency does not invite volunteers who may be willing to be updated in instrument landing procedures for wide-bodied aircraft at Lindbergh Field to stop by on their own time, for wine, cheese, and training. The technology works and there is too much at stake for its adoption. The hard parts of the process here are required participation (read "force") and the probable, if eventual, necessity that the change be initiated or induced, not simply reinforced or continued. (Mann, 1986, pp. 50-51)

According to Mann (1986), interventions with this strategy are "more likely to be high tech than high touch; they will emphasize product over process" (p. 50). Successful change is therefore signified by a finished product or the completion of goals. Critiques of this approach claim that the instrumental-rational approach to change fails to take into account teachers' beliefs, values, and ideas about education and learning.

A well known educational change model based on an instrumental-rational point of view is the Research, Development, and Diffusion (RD&D) Model. A typical example

of an RD&D staff development initiative is passing legislation to improve the quality of teachers or to raise standards. The principal assumption of this model is that successful, beneficial innovations can be mandated into use. As a result, change initiatives based on the RD&D model tend to rely on implementation strategies involving administered and legislated change (Havelock, 1971).

The RD&D model also emphasizes the systematic and sequential nature of knowledge creation and utilization and the possibility of developing “high performance” products that guarantee “user-proof” implementation. Basic tenets of the RD&D model assumes the existence of:

1. a rational sequence—research, development, packaging, dissemination—for evolving and applying a new practice
2. large-scale, lengthy planning
3. a division and coordination of labor directly related to the rational sequence and planning
4. a passive but rational consumer who accepts and adopts the innovation

(Havelock, 1971).

Finally, this model envisions change as an orderly, planned sequence beginning with problem identification, finding or producing a solution, and diffusing the solution. All efforts focus on research, development, and diffusion, with little attention to helping users implement the innovation.

According to Tye and Tye (1984), the instrumental-rational approach to change has encountered little success. Failure to account for people’s values, beliefs, skills, and

the culture to which they belong virtually guarantees that those involved will refuse to change or revert back to old behavior once the pressure to change has been removed. As a result, the majority of change models presented in this section and in the literature correspond more closely to the normative-cultural and personal-transformative approaches to change.

Normative-cultural educational change models

In the normative-cultural perspective, change is understood in terms of both personal dimensions and social contexts. Change models using this approach view change as a process affecting teachers' attitudes, values, skills, and relationships, while at the same time affecting and being affected by organizational contexts and relationships among people (Montgomery & Way, 1995).

According to Elmore (1997), educational change will not encounter widespread success until the general public ceases to regard good teaching as an individual trait (like shoe size or hair color) rather than as a professional norm, or a set of learned professional competencies. As long as good teaching is regarded in this manner, he maintains, it makes little sense to ask the broader systemic question about why more evidence of engaging teaching does not exist. Furthermore, Elmore adds that only 25% of teachers are intrinsically motivated to invest large amounts of their own time learning new ways to think about teaching, looking to outside models to improve teaching and learning, getting others to cooperate in changing their practice, and seeing their own practice in a broader social context. In addition, the remaining 75% of teachers are embedded in institutional structures that provide them with incentives to act the way they do.

In fact, Elmore (1997) maintains that the issue of incentives is crucial to achieving educational change on a large scale. Without external normative structures, which would give visibility and status to those who exemplify them, teachers have no incentive to think of their practice as anything more than a bundle of traits. Therefore, Elmore suggests the creation of strong professional and normative structures for good teaching practice that are external to individual teachers and their immediate working environment, and provide a basis for evaluating how many teachers are approximating good practice at what level of competence. External incentive structures recommended by Elmore include performance standards developed by professional bodies, alternative credentialing systems, salary increments for staff development related to changes in practice, release time to work on curriculum or performance standards, time to develop curriculum units that embody particular approaches to teaching, and opportunities to engage in demonstration teaching.

Likewise, Elmore (1997) is a proponent of small schools and other means of organizing schools into smaller sub-units, asserting that these structures create stronger group norms inside larger schools and provide more opportunities for adventurous teachers to connect with their less ambitious and reflective colleagues. According to Elmore, change is more likely to occur in these smaller teaching and learning environments where intrinsic motivation to engage in challenging practice is intensified and focused, as teachers participate in face-to-face relationships, interact around common problems of practice, and focus on students. In this way, teachers have the opportunity to exercise real influence over each other's practice.

Finally, Elmore (1997) maintains that plans for large scale educational change must include *intentional* processes for getting exemplary practices to scale; it is not enough to simply introduce a change, train a few teachers, and hope that the change will take off on a large scale. In fact, Elmore presents five methods for “getting to scale,” without showing a preference for any one approach or providing evidence of the success of any of them:

- Incremental growth: Train a given number of teachers/year, until 100% are teaching a particular way
- Cumulative growth: Create interventions that expose teachers to new practices, monitor the effects, and design processes to compensate for weaknesses in initial effects
- Discontinuous growth: An initial group learns a new practice, and members of each group work with another group, and so on (analogous to a chain letter)
- Unbalanced growth: Create “pioneer” schools of exemplary, high performing teachers that would, over time, be staffed with less accomplished practitioners who would be socialized into the norms of good practice
- Cell division or reproduction: Create settings where exemplary practitioners are concentrated and allowed to develop new approaches to teaching practice. Then, on a predictable schedule, these practitioners are asked to form a new school, using the genetic material of their own knowledge and understanding to recruit a new cadre of teachers that they educate to the new set of expectations about practice (Elmore, 1997, pp. 20-24).

Perhaps the most representative change model of the normative-cultural tradition was proposed by Fullan (1994), who stated:

“The change process is exceedingly complex as one realizes that it is the combination of individuals and societal agencies that make a difference” (p. 41).

A principal tenet of Fullan’s view of change is that education’s moral purpose is to make a difference in all student’s lives and to produce citizens capable of functioning productively in dynamically complex societies (Fullan, 1993). With the goal of preparing students for a continually changing world, Fullan sees teachers as in the business of continuous innovation and change. The “new problem with change,” however, is transforming schools into learning organizations that expertly deal with change as a normal way of life.

One of the difficulties in achieving change, according to Fullan (1999), is that “learning occurs on the edge of chaos” where a delicate balance exists between structure and anarchy, comfort and anxiety, collaboration and dissonance, and connectedness and incoherence. Indeed, he asserts that the nature of change at all levels of society is chaotic. While he does state that collegiality and high self-efficacy and self-esteem make change “work” for teachers (Fullan, 1991), Fullan (1994) is clear that “productive change is full of paradoxes, and components that are often not seen as going together” (p. 4). Like Stacey (1992), Fullan proposes that the change process is uncontrollably complex and, in many instances, “unknowable.”

Offering guidelines for understanding and acting in complex change situations, Fullan (1999) recommends that those involved in and organizing change initiatives recognize that conflict and diversity are “friends” of change, as they are commonly

associated with creative breakthroughs and learning from dissonance. In addition, Fullan urges awareness that change produces anxiety, which is best contained by emotionally intelligent individuals with strong ego structures and by collaborative cultures that offer emotional support to their members. Stating that the natural state of complex societies is confusion, Fullan also asserts that coherence is not reached through top-down or bottom-up strategies (the top is too distant and the bottom is too overwhelmed). Rather, middle-level administrators (i.e., principals) serve an essential meaning-making role as integrators and synthesizers.

Finally, Fullan maintains that there is no silver bullet of change. While theories of change can guide thinking and action, each situation and the context in which it exists are unique, making change unpredictable. In fact, he agrees with Stacey (1992), who states: “links between specific actions and specific outcomes become lost in the detail of what happens” (p. 20). Consequently, Fullan asserts that no definitive theory of change applies universally, and it is up to those involved in change to craft their own theories of change in their own situations.

Wertheimer and Zinga (1997) support Fullan’s notion that change is non-linear and dynamic. However, they allege:

“We further assert that, although much of the behavior [in a school reform project] appears to be random, chaotic, and unpredictable, within the randomness there exist patterns—points of stable attraction. Whether behaviors gravitate towards these points of attraction is dependent upon initial conditions in the organization” (p. 4).

Wertheimer and Zinga utilize the language and assumptions of Chaos Theory to describe change in educational organizations. They subsequently apply the following key components of Chaos Theory to their model of school reform:

- When a catalyst is applied to a system, *noise*—local disorder, turbulence, and fluctuations—becomes apparent. While this noise may appear random, it actually contains patterns that contribute to an understanding of a system’s complexity.
- The noise generated when a catalyst is applied to complex systems depends on *initial conditions*. Initial conditions in schools refer to school culture, such as organizational climate, norms, and values, and personal needs of school staff.
- Noise patterns are similar at all levels—individual, school, or district.
- The system is *deterministic*, i.e., it can be exactly predicted given information about initial conditions (Wertheimer & Zinga, 1997, p. 4).

Wertheimer and Zinga (1997) describe the change process as a recursive function that should make it possible to trace patterns back to initial conditions present in individuals and the environment. They purport that this function contains cultural variables in an organization, personal variables internal to individuals, and parameters that determine the influence of each variable. Cultural variables within an organization include the language used within an organization, peer pressure, the organizational hierarchy, and the use/effect of stereotypes on organizational behavior. Internal variables consist of individuals’ needs to control their environment (e.g., the classroom), the need to limit threats to one’s ego and self-esteem, the degree to which a person is flexible, and one’s

previous experience with change, innovation, and experimentation. The “initial conditions” in an organization are the initial values of its cultural and personal variables.

As with the models proposed by Fullan (1994) and Wertheimer and Zinga (1997), the Organizational Development (OD) model of educational change emphasizes group and organizational features that influence the degree to which an educational change is implemented. As such, it fits neatly into the normative-cultural approach to change. However, the OD model differs from the previous two models in that it focuses on group and organizational aspects of change to the exclusion of individual concerns. In fact, Schmuck, Runkel, Arends, and Arends (1971) described the OD model as leaving “...the processes and structures within individuals to the individuals themselves, their ministers, and their psychiatrists” (p. 9).

The basic assumption of the OD model of educational change is that groups and organizations are the major sources of problems related to changing schools. As a result, the primary goal of the OD model is to develop organizational adaptability, improve the effectiveness of the group or organization’s subsystems, and develop group members’ interpersonal skills (Hall & Hord, 1987).

Personal-transformative educational change models

In the personal-transformative approach to change, emphasis is on the personal nature of change and the transformation of teachers’ beliefs and practices. Change models utilizing personal-transformative approaches are based on two tenets: (1) behavior is based on beliefs and values; and (2) individuals—not groups or organizations—change. In this model, the personal nature of change threatens existing

interests and routines, while at the same time heightening uncertainty and complexity. As such, the personal-transformative approach to change involves examining personal conceptions, engaging in dialogue about these conceptions, justifying and reflecting on one's own practice, and, hopefully, transforming one's beliefs and practices (Montgomery & Way, 1995).

According to Peca (1994), the essential component of change is "giving up the old for the new" (p. 7). As such, Peca (1994) has suggested that educational change is accomplished through the grief process (Kubler-Ross, 1969), which is comprised of five stages: denial, anger, bargaining, depression, and acceptance. In the first stage of change, denial, resistance is strong as teachers might tend to deny that a change is going to occur. This occurs in response to the fact that their usual sense of security and comfort are threatened. In the anger stage, the reality of the change becomes persistent, and teachers "flail externally against the source(s) of change in hope that such expression of anger will cause the new reality to cease its progression" (p. 8). Through bargaining, teachers make a final effort to hold onto their old ways of operating by finding ways to fit the change into their existing repertoire of teaching behaviors. The fourth step in the grief/change process is depression. In this stage, teachers realize that change is inevitable and mourn the loss of old behaviors. Finally, teachers move ahead with new behaviors in the acceptance stage.

In contrast, the Social Interaction Model of educational change (Havelock, 1971), presents educational change as a series of decision phases through which individuals must pass as an innovation is diffused over time. These decision phases are: (1)

awareness of the innovation; (2) interest in and search for more information about an innovation; (3) decision to adopt the innovation; (4) trial; and (5) adoption of the innovation. Outright rejection of the innovation can occur at any time in this 5-step process. Interestingly, Havelock (1971) asserts that the role of the change agent is hardly needed after phase 1 (awareness of the adoption); a point that is contested by most in the staff development field.

According to Joyce and Showers (1983), change is a personal process in which teachers struggle, to varying degrees, to assume control of new practices that may differ from those they typically use in their classrooms. A 2-step process presented by Joyce and Showers (1983) stipulates that teachers need to go through two stages before “executive control,” or successful change can be considered to have occurred. According to Joyce and Showers (1983):

Executive control consists of understanding the purpose and rationale of the skill [being learned or acquired] and knowing how to adapt it to students, apply it to subject matter, modify or create instructional materials attendant to its use, organize students to use it, and blend it with other instructional approaches to develop a smooth and powerful whole (p. 8).

In this model, teachers experience stages of horizontal and vertical transfer prior to achieving executive control. According to Joyce and Showers (1983), horizontal transfer refers to a condition in which a skill can be shifted directly from a training situation to the classroom. On the other hand, vertical transfer refers to conditions in which a new skill cannot be transferred to the classroom until it is adapted to fit the specificities of a teacher’s individual situation or additional learning occurs. Until executive control is realized, the use of a new skill can be “confusing” and “laborious”

(Joyce & Showers, 1983, p. 9). An important factor, therefore, determining the speed at which executive control is achieved is the degree to which the new skill disrupts existing patterns of performance. A skill that is markedly different from a teacher's current repertoire will cause a great deal of discomfort and, possibly, avoidance.

While the above models of education change have garnered some acceptance, perhaps the most well-known and widely accepted models of educational change in the literature are Doyle and Ponder's "practicality ethic" (1977) and Hall and Loucks' Concerns Based Adoption Model (CBAM) (1976).

In 1977, Doyle and Ponder presented a model of educational change based on what they termed the "practicality ethic." Underlying this model is the view that an effective change strategy must be based on an understanding of the naturally existing mechanisms which operate in school environments. According to Doyle and Ponder, ecological variables shape the way teachers think about and conduct their work. Unfortunately, many procedural recommendations for the classroom simply lack ecological validity. That is, many proposed practices fail to mesh with existing features established by the structure and flow of real environments.

In particular, Doyle and Ponder (1977) assert that change strategies must focus on the decision making processes which underlie teacher reaction to change proposals. Messages that are seen as "practical" (on a variety of levels) will be incorporated into teacher plans. Therefore, the ultimate fate of an innovation depends on user decisions . This is the practicality ethic.

The practicality ethic is manifest in the common teacher practice of labeling change proposals with the term "practical." This term is an expression of teacher perceptions of the potential consequences of attempting to implement a change in the classroom. Recommendations that are seen as practical are ones which teachers will most likely try to incorporate into their classroom practices. Those that are seen as impractical have little chance of being tried unless control mechanisms are strong.

Teachers use three general criteria to determine if a change proposal is practical: instrumentality, congruence, and cost. Instrumentality refers to the degree to which a change proposal is perceived as immediately useful. Instrumentality is greatly increased if abstract principles and outcomes specifications are converted into procedural specifications. Teachers will find a change practical if they are shown how to do something rather than listen to inspirational or theoretical discussions of the rationale for or protected outcomes of the innovation.

Congruence refers to the extent to which a proposed procedure is congruent with teachers perceptions of their own situations. With regard to the implementation of innovations, congruence is the perceived match between a change proposal and prevailing conditions. Practices which depart radically from conditions which are normal for the teacher are usually viewed as impractical. A second aspect of congruence involves perceptions of the origins of the innovation proposal and the spokesperson for the innovation. Innovations targeted for use in a particular school setting have a stronger probability of adoption when they originate within that setting or when they have been developed in a highly similar setting. As Doyle and Ponder explain:

"A practice, for example, which is known to work in an upper-middle-class suburban high school may often be perceived as impractical by teachers in an inner-city school, especially when communicated by a university consultant" (p. 22).

A final element in congruence is role congruence, or the compatibility of the proposed change with teachers' self-image and preferred mode of relating to students. For example, Doyle and Ponder point out that career education projects often encounter strong resistance among secondary school teachers of academic subjects who do not view themselves as teachers of career development.

The third dimension of the practicality ethic is cost. According to Doyle and Ponder, "cost is the ratio between the amount of investment required to implement an innovation and return that may be realized" (p. 24). It refers to perceptions of the ease with which a procedure can be implemented and the potential return for adopting the innovation. The cost dimension involves the judgments about personal effort, social cost, and social reward. Costs in terms of personal effort are related to the complexity of the innovation. Social costs are determined by the reactions of students, colleagues, and administrators. In addition, Doyle and Ponder report that nonexistent or inconsistent reward systems in schools contribute heavily to teacher resistance to complex innovations. They add that teachers are especially responsive to social rewards such as recognition by administrators and colleagues and student enthusiasm.

Guskey (1986), on the other hand, presents an alternative model of teacher change, asserting that teachers will not change their beliefs and attitudes until they have changed their practices and gained evidence of changes in student outcomes. According

to Guskey, this model “is predicated on the idea that change is a learning process for teachers that is developmental and primarily experientially based” (p. 7). Only when teachers see that a new program or innovation enhances student learning will significant change in their beliefs and attitudes likely occur. The key to the success of this model, therefore, is convincing teachers to experiment with new teaching strategies and techniques at a point in time when they are not necessarily convinced of the need for or utility of change.

The Concerns-Based Adoption Model (CBAM), which was developed by Hall and Loucks in 1976, is a very popular educational change theory, not only for power in explaining educational change, but also for its utility in planning and evaluating staff development efforts. The CBAM model was developed by researchers at the Texas Research and Development Center who were conducting research on innovation adoption in educational institutions. During the course of this research, they became aware that people involved in the change experienced a series of *concerns* as they became more familiar with and adept in the use of an innovation. The creators of the CBAM model defined a concern as follows:

“The composite representation of the feelings, preoccupation, thought, and consideration given to a particular issue or task is called *concern*...All in all, the mental activity composed of questioning, analyzing, and re-analyzing, considering alternative actions and reactions, and anticipating consequences is a *concern*. An aroused state of personal feelings and thought about a demand as it is perceived is *concern*” (Hall and Hord, 1987, pp. 58-59).

Early concerns in the change process are with the self, followed by task-related concerns, and, finally, concerns about the impact of the innovations on others. According

to Hall and Hord (1987), it is necessary for early stage concerns to either be resolved or at least reduced in intensity before later, more mature concerns, can emerge or increase in intensity.

These concerns, as outlined in the Concerns-Based Adoption Model are briefly described below:

0: AWARENESS: Little concern about or involvement with the innovation is indicated.

1: INFORMATIONAL: A general awareness of the innovation and interest in learning more detail about it is indicated. The person seems unworried about him/herself in relation to the innovation.

2: PERSONAL: The individual is uncertain about the demands of the innovation, his/her inadequacy in meeting those demands, and his/her role with the innovation.

3: MANAGEMENT: Attention is focused on the processes and tasks of using the innovation and the best use of information and resources. Issues relating to efficiency, organizing, managing, scheduling, and time demands are utmost.

4: CONSEQUENCE: Attention focuses on the impact of the innovation on students in the teacher's immediate sphere of influence. The focus is on the relevance of the innovation for students, evaluation of student outcomes, including performance and competencies, and changes needed to increase student outcomes.

5: COLLABORATION: The focus is on coordination and cooperation with others regarding the use of the innovation.

6: REFOCUSING: The focus is on exploration of more universal benefits from using the innovation, including the possibility of major changes or replacement with a more powerful alternative (Hall & Hord, 1987, p.60)

A major assumption of the CBAM model is that change has an important personal side, and that “without understanding where the clients ‘are,’ only through chance will interventions made by change facilitators address the needs of innovation users and non-users (Hall & Hord, 1987, p. 8). The model also assumes that attention must be given to individuals and their use/non-use of an innovation, as well as their individual experiences with the change process. Only when an understanding of the individual is achieved can the change process for an entire district or school be examined (Hall & Hord, 1987).

In addition to identifying these stages in the change process, the CBAM model includes practical techniques enabling change agents to determine at which stages in the change process participants are located, plan professional development strategies that are appropriate to each stage of concern, and to track individuals progress through the change process over time.

A Brief History of Staff Development for Teachers in the United States

In their examination of the historical forces that developed and characterized staff development for teachers in the United States, Edelfelt and Lawrence (1975) identified twelve features that they believed guided staff development through the mid-1970s. As such, they are essential for understanding the history of in-service education and the forces that continue to face staff developers and change agents:

1. The primary role of the school is the giving and receiving of information.

2. Learning is the receiving of information to be stored and used later.
3. Curriculum and teaching are relatively fixed elements in the school.
4. The main business of teacher education is the quest for mastery of some relatively stable subject matters and methods of teaching.
5. In-service education is training that is designed, planned, and conducted for the teacher by persons in authority.
6. The central purpose of in-service education is the remediation of teachers' deficiencies in subject matter.
7. Leadership is "direction from above," and motivation is "direction from outside."
8. Supervision is diagnosis, prescription, modeling, inspection, and rating.
9. Teacher education in teacher preparation institutions and teacher education in schools are separate and discontinuous processes.
10. Intellectual leadership in goal setting and planning for in-service education appropriately comes from outside the school.
11. The teacher is a solo practitioner.
12. Prescriptive legislation is an appropriate vehicle for improving the quality of teaching standards (Edelfelt & Lawrence, 1975, p. 1).

In 1975, Edelfelt and Lawrence asserted that the above 12 concepts were all "fading and being replaced" (p. 9) at different rates. In the year 2000, however, vestiges of these concepts can still be felt and must be recognized in order to fully understand how and why staff development has taken the turns that it has.

Programs for the professional development of teachers in the United States have a short history. In fact, most 19th century teachers received no formal preparation for teaching at all. Rather, teacher selection was based primarily on judgments of moral character, with most teachers completing little more than an elementary education. Because school was generally in session for a few hours per day for two or three months of the year, most schools dealt with a narrow range of information and teacher training was neither required nor available (Edelfelt & Lawrence, 1975).

The first formal programs for the professional preparation and development of teachers began in the 19th century under the leadership of advocates such as Horace Mann, James Carter, Calvin Stowe, and Thomas Gallaudet. The first professional school for the preparation of teachers, or “normal school,” was established in 1838. However, standards for admission were minimal, simply requiring that prospective teachers pass a basic exam, be of good moral character, and declare an intent to teach (Bernier & McClelland, 1989). At the beginning, the purpose of teacher institutes and normal schools was to expand teachers’ knowledge of the subjects s/he was teaching. Later, training began to include principles of teaching. At this time, there was a strong tendency for teachers “to seek security in a quest for *the* teaching techniques” (Edelfelt & Lawrence, 1975, p. 11). Furthermore, the principal training methods used in teacher institutes and normal schools were lecture and recitation, techniques that teachers were then expected to use with their students.

After the Civil War, collegiate institutions began to offer courses in pedagogy and, toward the end of this century, teachers' colleges appeared. At this point, practice

teaching became part of most preservice programs. It was not until the end of World War I, however, that the two and three year normal or teacher training schools were by and large replaced by 4 year teacher preparation programs granting bachelor's degrees (Bernier & McClelland, 1989). In fact, by 1930, 75% of all teachers had attended two or more years of college (Edelfelt and Lawrence, 1975).

Until this time, the main purpose of staff development had been viewed as remediating teachers' deficiencies in subject matter. However, the child-study and progressive education movements of this era shifted the emphasis of staff development to other aspects of teachers' practice. In fact, the term "workshop" was introduced in the 1930s to describe "problem-solving, action-oriented in-service workgroups" (Edelfelt & Lawrence, 1975, p. 10). According to Edelfelt and Lawrence (1975), however, lecture-discussion remained the dominant form of learning.

After World War I, strong emphasis was placed on legislated standards to improve the quality of teachers. In fact, attendance at teacher institutes was compulsory in most of the United States. Furthermore, when the supply of potential teachers exceeded demand during the Depression years, standards for teaching were improved. Furthermore, an unprecedented era of curriculum reform was initiated immediately following World War II. Reformers, who sought to improve schools through the enhancement of curriculum materials, solicited the participation of the academic community in their initiatives, to the exclusion of classroom teachers. At this time (as in the later Sputnik era), staff development shifted back to the notion of remediating teachers' deficiencies in their knowledge of specific subject matter (Edelfelt & Lawrence,

1975). On the whole, these curriculum reforms were not successful. However, they did result in a proliferation of curriculum materials, and they drew attention to the notion of school improvement.

Staff development began to take hold in the 1960s. More laws were enacted to raise standards, causing many teachers to begin accumulating college and other staff development courses, many of which were not relevant to their performance as teachers (Edelfelt & Lawrence, 1975, p. 14). Staff development, primarily in the form of a menu of short-term workshops, was structured around incentives for advancement and salary enhancement and on policies requiring the renewal of teaching licenses via the accumulation of continuing education credits (United States..., 1998).

In response to Sputnik, the federal government spent billions of dollars on staff development from the 1960s to the 1970s, mainly as part of the National Defense Education Act and National Science Foundation Institutes. According to Edelfelt and Lawrence (1975), the purpose of such staff development was generally to improve teachers' subject matter knowledge or orient teachers to a new curriculum. Typically, a single teacher per school district was selected to participate in a summer, semester, or year-long training course. Once back in the classroom, however, these lone teachers were not equipped to change entire schools.

In this era, staff development situated in school improvement efforts tended to be “mandated, prescriptive, remedial, content-focused, and organized and implemented by persons in authority” (Holly, 1989). Edelfelt (1983) points out that the assumption at this time was that improving individual teachers would improve schooling and, consequently,

students' learning. Most change efforts, relying on the "research-development-diffusion" approach, were "top down" approaches designed to implement programs that "experts" had developed. Teacher participation in reform initiatives was not sought. Indeed, reformers strove to develop "teacher-proof" curricula.

In addition, staff development initiatives in the 1960s tended to address one aspect of an issue while ignoring others. For example, new curriculum materials were frequently introduced to teachers while ignoring the skills teachers might need to implement them (Wideen, 1987). However, implementation studies of the 1960s staff development initiatives revealed that the majority of these staff development initiatives had *failed* (Fullan, 1994). According to Fullan (1994), "there followed a period of stagnation, recovery and regrouping during most of the 1970s" (p. 1). In response to these disappointments, some new ideas were brought to the forefront; perhaps the most striking of these new ideas was that teachers must play a vital role if school improvement is to be successful and that a variety of change strategies should be employed in a given situation. By the end of the decade, the effective schools movement also began to garner evidence that schools can make a difference in a variety of conditions.

While staff development in the 1970s still tended to focus on the delivery of workshops and training, its focus did broaden from an emphasis on the remediation of individual teachers to an emerging recognition of the need to focus on the entire organization and the individuals who comprise it--school-based staff development (Killion & Harrison, 1997). According to Hyde and Pink (1992), the birth of the National Staff Development Council in the late 1970s indicated that a growing number of

educators advocated a conception of continuing professional growth for practicing teachers that emphasized issues related to decision making, process, and professional knowledge (Blackman, 1989), in contrast to the traditional notion that teachers needed staff development to remediate their inadequacies. Lessons learned in this decade also included the realization that restructuring could not be imposed from the top-down and that the involvement of teachers was vital to programmatic success (Fitch & Kopp, 1990). This signaled the beginning of a trend toward *proactive* visions of professional development rather than narrow, *reactive* ones.

According to Sparks and Loucks-Horsley (1990), staff development “came of age” in the 1980s, with numerous conferences, workshops, articles, books, and research reports on the subject. Furthermore, state legislatures and local school districts had come to view staff development as a key aspect of school improvement efforts. This was due in large part to the lessons learned in the previous decades, as well as to some significant events in the 1980s, such as the 1983 publication of *A Nation at Risk*. By the mid-1980s, the focus of staff development reflected a movement toward organizational development, school improvement, and systemic change. The role of the staff developer also changed, from expert trainer to collaborator in problem solving, facilitator, fact finder, technical specialist, data collector, reflector, and more (Fitch & Kopp, 1990; Killion & Harrison, 1997).

The last two decades have been characterized by a near constant focus on school reform. As opposed to the previous era, single individual innovations are not being considered one at a time; rather, comprehensive reforms are customary (Fullan, 1994).

The roots of school reform in the 1980s and 1990s lie not only in a generalized concern with schools' deficiencies, but also in concerns related to American competitiveness in a global economy, the need to address demographic shifts, an "unraveling of the social fabric," cynicism about the government, and changing political philosophies (Lashway, 1999; Fenstermacher & Berliner, 1985). The focal points of the present "reform era" (1996-present) are standards, accountability, and privatization.

Furthermore, five "streams" of reform characterize the present era:

- *Reforms in subject matter standards, curriculum, and pedagogy* aspire toward more ambitious student outcomes and require a fundamental change in the nature of students' intellectual tasks and teacher-student relations
- *Reforms centered on problems of equity among a diverse student population* require teachers to identify and alter classroom practices that contribute to student failure and that undermine equal opportunity to learn
- *Reforms in the nature, extent, and uses of student assessment* require widespread, rigorous use of authentic assessment on the part of teachers
- *Reforms in the social organization of schooling* necessitate broad school restructuring oriented toward new principles
- *Reforms in the professionalization of teaching* center on teachers' demonstrated knowledge base, on conditions surrounding teacher certification and licensure, and on the structure of career opportunities in teaching (Little, 1993, pp. 130-132)

In the present era (1990-present), there is widespread agreement that staff development for teachers is crucial to school improvement. The reform agenda requires

most teachers to reconceptualize their practice, to construct new classroom roles and expectations about student outcomes, and to teach in ways they have never taught before and probably have never experienced (Nelson & Hammerman, 1996). Therefore, the professional development of teachers is of primary importance if the reforms are to achieve the improvements sought. Teachers must learn new skills and perspectives and unlearn practices and beliefs about students or instruction that had dominated their entire professional lives. In fact, many today view the professional development of teachers as the central issue that can make standards-based reform work (United States..., 1998).

Others consider staff development crucial in this area, but for reasons other than school reform. For instance, Wideen (1987) stipulates that continuous staff development is necessary to keep up with the enormously increasing knowledge base for which teachers are responsible. In addition, Wideen reports that recent ethnographic studies of classes in schools have produced a new awareness of the instructional process. Furthermore, the developmental stages through which teachers progress during their careers and while implementing an innovation have been examined in depth, producing new models from which to choose when planning programs of staff development. The last three decades have shown that even a modest innovation requires changes of a complex nature, and what has been learned about the school as a social structure indicates that teachers working in isolation will be unlikely to have any effect on school improvement. Finally, Wideen recognized that even the very best of pre-service teacher education cannot equip one for a lifelong career. In light of an expanded knowledge base

and the continuing rate of change that is occurring in society, the need for continued professional growth among teachers is vital (Wideen, 1987).

Dominant 20th Century Staff Development Models

In 1992, Hyde and Pink wrote:

“As we examine the literature on staff development, we are struck by its essentially atheoretical nature and how often conceptions are informed more by personal experience than by systematic reflection and research...Rarely is any well-constructed theoretical orientation present (pp. 4-5).”

As suggested above, few theoretical models of staff development are found in the literature. Further confusing this issue is the fact that the term “model” is used in the staff development literature to describe a variety of different concepts. On the one hand, some researchers employ the term “model” to explain the perspectives (e.g., social, cultural, political) from which to view and understand staff development for teachers. On the other hand, most of the staff development literature utilizes the term to describe specific staff development strategies and techniques, the role(s) of staff developers, and ways to organize staff development objectives.

Staff development “models” as perspectives from which to view and understand staff development

The most widely known theoretical models of staff development were postulated by McLaughlin and Berman (1977) and others. McGlaughlin and Berman contrast staff development initiatives based on a *deficit model* to those utilizing a *developmental model*. According to the deficit model, schools’ and teachers’ problems are caused by inadequate information and inadequate skills. In other words, something is lacking and needs to be

corrected. As a result, a principal tenet of the deficit model is that adequate information and skills will solve educational problems. Hence, deficit model staff development programs are “top-down.” “Experts” in positions of authority determine teachers’/schools’ needs and design staff development activities to correct deficiencies (McGlaughlin & Berman, 1977). This model is supported by the research of Borko and Putnam (1996) which suggested that weak content mastery on the part of teachers is related to lower student performance. Clearly, this staff development model is built on the Research, Development, & Diffusion (RD&D) model of educational change that dominated staff development initiatives in the 1960s and 1970s.

Blackman (1989) posits a model of staff development that is based on both the social and historical contexts in which staff development activities occur and on how one views the roles of the persons involved. Blackman posits that social and historical trends in the United States in the last century have produced different visions of the role of the classroom teacher, including craftsperson, technician, professional, etc. According to Blackman, if the teacher is viewed as the applier of a craft, then staff development will focus primarily upon the methods and techniques of teaching. If teachers are viewed as technicians functioning in isolation from one another, staff development will focus upon the activities of the classroom. If the teacher is viewed as a professional, staff development will address issues related to decision making, practice, and professional knowledge (Blackman, 1989).

On the other hand, the developmental model of staff development represents an opposing staff development model which is based on “a set of expectations about the role

of teachers, about their professional needs, and about their responsibility for solving their own problems in the classroom” (McGlaughlin & Berman, 1977, p. 193). In this model, emphasis is placed on equipping teachers with the resources and authority to perform their jobs, offering them a variety of staff development options, a flexible program format, and concrete, ongoing, individual and small-group learning opportunities (McGlaughlin & Berman, 1977). Typical staff development activities include: teacher study groups, curriculum writing groups, action research, internal program evaluations, and peer-organized workshops and seminars (Guskey & Huberman, 1995).

Staff development “models” as specific staff development strategies and techniques

More often than not, the term “model” is used in the staff development literature to describe a set of staff development strategies and techniques, staff development objectives, and/or staff developer roles. After a careful review of the work of Sparks and Loucks-Horsley (1990), Orlich (1988), Gall and Vojtek (1994), Killian and Harrison (1997), Darling-Hammond (1995; 1996; 1998) and others, it is possible to conclude that eight staff development models dominate the staff development literature. Each model describes different ways teachers can engage in professional growth. According to Gall and Vojtek (1994), each model is effective for a different purpose. They even go so far as to say that the models can be located on a continuum of simple to complex. However, anyone who studies these models will soon observe that these models are in no way mutually exclusive; the techniques and strategies used in the various models frequently overlap. These models, which are described in the section that follows, are: the training model, the individually guided staff development model, the observation/assessment

model, the action research model, the organizational development model, the development/improvement process model, the professional development school model, and the organizational context model.

Training Model

For many people, training is synonymous with staff development. Historically, training has also been considered a cost-efficient way for teachers to acquire knowledge skills. This model includes the workshop-type sessions in which the presenter is the expert who establishes the content and flow of activities. Typically, the training session is conducted with a clear set of objectives or learner outcomes which may include awareness, knowledge, skill development, changes in attitude, and/or transfer of training. Training activities may include exploration of theory, demonstration or modeling of a skill, practice of a new skill under simulated conditions, feedback about performance, and follow-up coaching or peer observation in the classroom. The primary role of the staff developer in this model is that of trainer/designer, one who designs and delivers learning experiences that help others acquire new skills, knowledge, and attitudes (Killion & Harrison, 1997, p. 2).

An underlying assumption of the training model is there are behaviors and techniques worthy of replication by teachers in the classroom. Another assumption is that teachers can change their behaviors and learn to replicate behaviors that were not previously in the repertoire (Sparks & Loucks-Horsley, 1990).

Research on the Training Model

Research on training since the 1980s has revealed that training alone (unaccompanied by practice or follow-up assistance) is highly effective at changing teacher knowledge, but less effective at achieving transfer of new teaching practices to the classroom or influencing student achievement. For example, five experimental studies of teaching effectiveness (Crawford, Gage, Corno, Stayrook, Mitman, Schunk, Stallings, Baskin, Harvey, Austin, Cronin, & Neuman, 1987; Anderson, Evertson, & Brophy, 1979; Stallings, Needels, & Stayrook, 1979; and Gage & Coladarci, 1980) tested the success of a wide range of instructional training strategies—from sending training packets to teachers by mail to holding generic, “canned,” training sessions to conducting several long, individualized workshops. With the exception of the 1980 Gage and Coladarci study, the principal conclusion of each of these experiments was that training brought about significant change on the part of teachers. For example, Crawford, et al. (1987) discovered that the difference in achievement scores between treatment and control classes yielded an effect size of .69, indicating that the mean score for classes of trained teachers corresponded to the 75th percentile rank of the classes of untrained teachers. Furthermore, interviews with study participants caused researchers to further speculate that specific assistance in implementing behavior changes would have perhaps led to increased training effectiveness in all studies (Crawford et al., 1987; Anderson, et al., 1979; Stallings, et al., 1979).

A meta-analysis of 91 staff development studies conducted from 1968 to 1983 (Wade, 1985) concluded that in-service training of any kind resulted in a mean effect size

of .52 for treatment groups (as opposed to control groups who do not participate in training). Furthermore, Wade (1985) identified the following effect sizes for training directed at each of the following levels of impact:

<u>Level of impact</u>	<u>Effect size</u>
teacher learning	.90 (highly effective)
teacher behavior	.60 (moderately effective)
teacher reactions	.42 (moderately effective)
student achievement	.37 (mildly effective)

(Wade, 1985, p. 50)

These findings suggest that training is more successful at impacting teacher concepts and knowledge (i.e., teacher learning) than anything else. An unexpected characteristic of these findings, however, is the fact that training was found to have a moderate effect on teacher behavior (.60) yet such a weak effect on student achievement (.37). This contradicted the accepted notion that changes in teacher behavior were essential to changes in student achievement.

In an attempt to identify the most effective components of staff development training initiatives that include components in addition to instruction, Joyce and Showers (1980) analyzed more than 200 studies in which researchers had investigated the effectiveness of various training methods. The major components of training in the studies Joyce and Showers reviewed were: (1) presentation of theory or description of skill or strategy; (2) modeling or demonstration of skills or models of teaching; (3) practice in simulated and classroom settings; (4) structured and open ended feedback in

the training setting; and (5) coaching (i.e., hands-on, in-classroom assistance with the transfer of skills and strategies to the classroom) (p. 380).

Joyce and Showers (1980) determined that the above training components were more effective in combination than on their own. For example, they found that simple presentation of theory did little more than impact teachers' awareness levels, while the presentation of theory, modeling of teaching strategies and opportunities to practice and receive feedback impacted teachers on three levels: awareness, concepts and organized knowledge, and principles and skills. For the impact of training to progress to the transfer level, Joyce and Showers asserted that training components must include presentation of theory, modeling, simulated practice and feedback, and *coaching for application*. Coaching for application includes direct coaching by peers, supervisors, consultants or other qualified individuals on how to apply new skills and models into teachers' day to day work, as well as making specific plans as to how to help students adapt to and benefit from a new approach. Unfortunately, Joyce and Showers' conclusions are weakened by the fact that they reported no effect sizes to back up their assertions of the effectiveness of the techniques they advocated. Yet, Joyce (1995) stated:

By the mid-1980s, we could assert that the art and science of training design had reached the point where designs could virtually assure that nearly all educators could master just about any teaching skill or model of teaching yet developed; follow-up activities by the educators themselves (peer coaching) could assure that nearly all of them would find satisfaction through implementation of the additions to their repertoire and the responses of the students (p. 13).

Following the introduction of Joyce and Showers' training model, several researchers conducted studies that either tested or were based on this model. Murphy (1987) reported the results of a staff development initiative aimed at increasing the learning of students and changing the workplace of teaching. Working with Joyce and Showers, she implemented a program consisting of presentation of theory, demonstrations of teaching models, expectations for practicing the models, peer coaching in classrooms, and study groups. According to Murphy, teachers and students in the 12 schools where teachers participated in the above training process achieved impressive student-level results, including fewer student suspensions, more promotions based on merit, and more reading and writing. However, Murphy (1987) did not account for possible threats to external validity, such as lack of random assignment, multiple-treatment interference, and reactive effects.

Similarly, a study designed by the Napa, CA County Office of Education (Robbins & Wolfe, 1987) to measure the effect of a Hunter-based staff development project on teacher behavior, student engagement rates, and student achievement revealed that Joyce and Showers' training model resulted in significant changes in teachers' appropriate use of instructional skills, significant ($p < .001$) improvement in student engagement, and improved reading and math test scores over a 3 year period. Decreases in all of these measures during the fourth year of the project were attributed to the cessation of coaching sessions; however, Robbins and Wolfe possessed no clear data indicating that this reversal in gains was truly associated with an insufficient amount of coaching.

In contrast to the findings of Joyce and Showers (1980), Wade's analysis of 225 cases involving coaching revealed no evidence that it increased training effectiveness. Similarly, Sparks (1983, 1986) analyzed the impact on teaching practice of three alternate training combinations: (1) presentation-demonstration-practice-feedback; (2) presentation-demonstration-practice-feedback plus 2 classroom observations by a peer; and (3) presentation-demonstration-practice-feedback plus 2 in-classroom coaching sessions provided by a trainer. Interestingly, Sparks found that peer observation training activities were more powerful than coaching or workshop-only activities with respect to the transfer of new teaching techniques to the classroom. However, Sparks' findings were weakened by a lack of random assignment and an extremely small sample size (N=19), both of which were openly acknowledged by Sparks. These findings support the work of Kerman (1979), whose experimental study compared the achievement of students in classes taught by teachers engaged in a staff development project consisting of training plus peer observation versus students in control classes. At the end of the 3-year study, Kerman found that approximately 2000 identified low achievers in experimental classes showed statistically significant academic gains over their counterparts in the control classes. Experimental students were also statistically significantly less likely to be absent and receive discipline referrals.

Likewise, Gliessman, Pugh, Dowden, and Hutchins (1988) reviewed 26 experimental studies to identify the effect of training method (instructional vs. practice-based) on skill changes through training. Limiting the subject of the studies analyzed to staff development efforts aimed at increasing teachers' questioning skills, Gliessman, et

al. found no significant difference ($z=.33$, $p>.05$) in the effect sizes for treatments involving only instruction (.79) versus practice-based treatments including overt practice of questioning skills and individual feedback (.84). Gleissman, et al. (1988), however, cautioned against generalizing from the results of this study, admitting that the conceptual basis of the 26 studies reviewed was weak and results based on a review of questioning skills might not be applicable to other teaching skills.

Many staff development researchers and practitioners have explored whether Joyce and Showers' training model or any other combination of training activities might ensure success in any setting. Interestingly, Wade (1985) also reported that there was no "magical combination of methods for instruction" (p. 52), contradicting Joyce and Showers' assertion that the presentation of theory- modeling-simulated practice and feedback- coaching for application combination produced better results than other techniques, either alone or in alternative combinations.

Furthermore, Wade's study concluded that the duration of training was insignificant. Staff development training lasting a few hours was found to be no more effective than training lasting more than 30 hours. Similarly, there was no significant difference in effect sizes between programs lasting less than six months and those lasting more than six months. These findings were corroborated by Gliessman, et al. (1988). However, these findings appear to "go against the grain" of most of what is accepted in the staff development field. For example, Sparks, Joyce, and Bennett's 1987 meta-analysis of over 200 staff development studies concluded that at least 25 episodes in which a new strategy is used are necessary for all of the conditions of transfer to be

achieved. In addition, Goldenberg and Gallimore (1991), Borko (1997) and others recommend that staff develop projects last one year or more.

Teachers need one to two years of intensive work to achieve a reasonable mastery of [new] skills...This insight from our research helps explain why so many past efforts to change teaching practices have not succeeded. Many of the needed changes in teaching are hard to achieve, and there is no tradition in U.S. education for training of the length and intensity required to achieve substantive change. (p. 72)

More research has been conducted on the training model than on any of the other staff development model combined. For the most part, there is agreement that training of any kind is effective at influencing teachers' attitudes and behavior. However, this is where consensus seems to stop. Various studies have produced contradictory evidence regarding the necessary methods and components of effective training efforts, as well as the recommended duration of training. As a general rule, evaluation reports and experimental studies have concluded that certain combinations and sequences of training activities are superior to instruction-only training. In addition, the results of these studies have recommended that adequate duration of training is essential to effective staff development. On the other hand, the findings of meta-analytic studies have revealed the opposite, concluding that instruction-only training is no more effective than instruction plus practice-based activities and that the duration of training does not influence its impact on teachers and students.

Individually Guided Staff Development Model

In the individually guided staff development, or self-instruction, model teachers design their own staff development, determine their own professional development goals,

and select the activities that will result in the achievement of these goals. An underlying assumption of this model is that individuals are capable of self-direction and self-initiated learning and that they can judge best their own learning needs. It also assumes that adults learn most efficiently when they initiate and plan their own learning activities and are most motivated when they select their own learning goals on the basis of a personal assessment of needs. In addition, this model corresponds to the Concerns-Based Adoption Model of change (Hall & Loucks, 1976), which states that, individuals experience different types of concerns as they learn new behaviors and change their practice, thereby necessitating individualized staff development activities.

According to Sparks and Loucks-Horseley (1990), the reading of professional publications and scholarly research, engagement in discussions with colleagues, attendance at a workshop or conference related to a teacher's concern, designing and carrying out special professional projects supported by incentive grants, and experimentation with new instructional techniques are all examples of individually guided staff development. In this model, therefore, staff developers are resource providers whose main function is to provide or link clients with resources that will help them reach their desired outcomes (Killion & Harrison, 1997, p. 5).

Research On Individually Guided Staff Development

On the whole, the research on individually guided staff development is very optimistic with regards to independent study and independent study groups. For instance, several studies support the use of self-instructional use by adult learners. Edwards (as

cited in Murphy & Lick, 1998), for example, concluded that teachers who use self-instructional materials learn as well as do those with supervision or instruction.

Independent study groups are initiated by teachers who decide that they have a common need and agree to get together to address that need. Participation is voluntary and not part of a larger school-wide design (Murphy & Lick, 1998). Lawrence (1974), who reviewed 97 studies and evaluation reports of staff development, concluded that self-instructional training units had a sound record of effectiveness. Wade's meta-analysis (1985) concluded that independent study produced the highest effect size of the structures (workshops, courses, mini-courses, institutes) she analyzed. She hypothesized that these results could be explained by the fact that the "most highly motivated people" tend to select independent study. Wade's position, however, is weakened by the fact that she neither identifies the mean effect size associated with independent study nor supports her assertion that independent study attracts the most motivated individuals.

Joyce, Murphy, Showers, and Murphy (1989) reported that the success of the study group design depends on two factors: the inclusion of a peer coaching component and whole school participation. According to these researchers, study groups implemented without peer coaching and by groups of enthusiastic volunteers achieve only a 10% transfer rate. When groups of volunteers participated in study groups plus peer coaching, transfer of content to classroom practice ranges from 75% to 90%. However, Joyce, et al. (1989) asserted that the transfer rate approaches 100% when entire faculties are involved and peer coaching is included.

In contrast, research on the outcomes of teachers' independent participation or attendance at workshops or conferences is less positive. Lawrence (1974) found that "single-shot" workshops that are not part of a general staff development plan were not likely to benefit teachers in terms of changes in teaching practice. As a further illustration, Bredeson and Scribner's (2000) study of 301 teachers who attended a 3-day professional conference consisting of workshops, round table discussions, work groups, consultation time with experts, and presentations by invited speakers revealed that, while teachers elected to attend the conference to gain an awareness of concepts, theories, how-to knowledge, and political knowledge from conference activities, only 30% planned to make changes in their classroom practices as a result of information obtained at the conference. Furthermore, few respondents were confident that they could disseminate newly acquired information at their schools. While the use of a convenience sample weakens Bredeson and Scribner's study, these results do suggest that individually guided staff development characterized by attendance at professional conferences is not likely to result in transfer to the classroom or, subsequently, changes in student achievement.

In summary, research on the individually guided staff development model is inconclusive. Studies have suggested that self-study and study groups may impact the practices of small pockets of teachers. However, research points to evidence that individually guided instruction is not likely to influence classroom or school-wide change without full staff participation or a school-wide support.

Observation/assessment model

The observation/assessment model is also known as the clinical supervision model (Gall & Vojtek, 1994). It can take many forms, such as peer coaching, clinical supervision, and teacher evaluation. The staff developer's principal responsibility is to help teachers transfer structured or unstructured learning experiences into practice. As such, coaches visit teachers' classrooms, observe teachers at work, gather information about teachers' behaviors, and offer feedback and support (Killion & Harrison, 1997, p. 4).

An assumption of this model is that observation and assessment of instruction provide the teacher with data that can be reflected upon and analyzed for the purpose of improving student learning. A second assumption is that reflection by an individual on his or her own practice can be enhanced by another's observations. A third assumption is that observation and assessment of teaching can benefit both the teacher being observed and the observer. The observer benefits by watching a colleague, preparing the feedback, and discussing the common experience. The teacher benefits by another's view of her or his behavior and by receiving helpful feedback from a colleague. A final assumption is that, when teachers see positive results from their efforts to change, they are more apt to continue to engage in improvement. Because this model can involve multiple observations and conferences over time, it is thought that the process can help teachers see that change is possible.

Research on the Observation/Assessment Model

In general, the effects of the observation/assessment model have not been studied at the level of classroom transfer or impact on students. Typical of the research on this model is a study conducted by Sparks and Bruder (1987) in which the effects of peer coaching were examined to determine whether peer coaching would improve collegiality among teachers, encourage experimentation with new practices, enhance teaching effectiveness, and cause teachers to react positively to the peer coaching process. Using survey and interview techniques, Sparks and Bruder determined that peer coaching produced helpful professional interaction, reduced teachers' fears of trying out new practices, and made teachers more comfortable with peer coaching. While 70% of teachers surveyed indicated that their students were "very likely learning more as a result of the skills [teachers] had received" (p. 56) from the peer coaching project and nine teachers backed up their claims with specific examples, Sparks and Bruder's study did not provide evidence that peer coaching impacted teaching strategies or student achievement.

Likewise, Busher (1994) examined the effects of peer coaching on a sample of 23 elementary teachers' professional attitudes and beliefs, compared to a similar sample of 26 teachers not involved in peer coaching. Using a causal-comparative research design, Busher was not able to reject the four null hypotheses of this study, namely that peer coaching would have no effect on teachers' professional attitudes and beliefs about teaching, level of comfort taking risks and trying new teaching techniques, feelings of isolation in the teaching environment, and feelings of job satisfaction. While Busher

herself pointed out that the generalization of these study results was limited due to small sample size, brief program duration, and the use of volunteers (who might naturally be more motivated and willing to experiment), this study provides an example of the superficial level of impact (attitudes/feelings) typically examined in studies of peer coaching and of the lack of significant effects associated with it.

An exception to the tendency not to focus on classroom transfer and student change in the observation/assessment research is a study conducted by Showers in 1983. In her first study of the effects of coaching teachers in the use of new instructional models, Showers (1983) explored two questions: whether teachers who received coaching applied these new models more effectively in the classroom than uncoached teachers and whether the use of new models impacted student learning. After conducting training for 17 teachers in the use of new models of instruction, Showers herself coached nine teachers, observing them once a week, providing feedback on teaching performance, furnishing encouragement and support, and assisting teachers in planning subsequent lessons. At the conclusion of the study, Showers concluded that coaching had definitely impacted both teacher implementation of new models and student achievement. Classroom observations revealed that coached teachers were better able to integrate new models into their day-to-day teaching, as opposed to their control group counterparts. Secondly, students of teachers who made extensive use of new models scored higher on an objective knowledge test than did students of teachers who did not use the models (Fielding and Schalock, 1985). Six to nine months after the conclusion of Showers 1983 study, Baker (1985) revisited the 17 study participants and found that coached teachers

still made more use of the instructional models in their classrooms than did uncoached teachers. However, he also found that even uncoached teachers had improved in their ability to use the new models since Showers' earlier observations, suggesting that uncoached teachers may be capable of making the same changes as coached teachers, albeit in a longer timeframe.

In conclusion, research on the observation/assessment model is fairly clear that this mode of staff development is effective at improving collegiality, promoting professionalism, and improving teachers' attitudes. In addition, its impact on contextual factors and teacher attitudes suggests potential linkages to the organizational context model of staff development. However, research regarding the effects of this observation/assessment model on classroom teaching practices and student achievement is deficient. Before substantive conclusions regarding the impact of this model can be made, it will be necessary to further study the effectiveness of this model.

Action research model

The action research model consists of teacher inquiry, or action research. In the action research model, individuals or groups identify a problem of interest. They then explore ways of collecting data on the problem. Data, which can range from examining existing theoretical and research literature to gathering original classroom or school data, are subsequently analyzed and interpreted by an individual or the group. Finally, changes in practice are made, and new data are gathered and analyzed to determine the effects of the change (Sparks & Loucks-Horsley, 1990).

An underlying assumption of the action research model is the belief that teachers can formulate valid questions about their own practice and pursue objective answers to those questions . Loucks-Horsley et al. (1987) also list three assumptions about teachers in the action research approach to staff development: (a) teachers are intelligent, inquiring individuals with legitimate expertise and important experience; (b) teachers are inclined to search for data to answer pressing questions and reflect on the data to formulate solutions; and (c) teachers develop new understandings as they contribute to and formulate their own questions and collect data to answer them.

Research on the Action Research Model

Typical research on the effectiveness of the action research model relies on anecdotal information and zealous enthusiasm about the potential impact of this form of staff development on teaching and learning. What's more, studies attempting to assess the impact of action research in a reliable fashion tend to focus on teacher-level attitude or organizational changes, rather than changes in classroom practice or student achievement. For instance, Rich (1983) asserted that action research increased teachers' willingness to change and motivated them to focus more on finding out what their students know. Furthermore, Lieberman (1986) concluded that action research was capable of increasing interaction among colleagues, increasing teachers' interest in applying research findings, and giving teachers a sense of empowerment.

A typical action research study is represented by the work of Burns (1995), who analyzed the impact of action research on 30 teachers in Australia. Burns reported that anecdotal information and teacher comments illustrate the utility of the action research

method in creating an intimate engagement with practice, increasing teachers' personal sense of growth, increasing teacher self-awareness and insight, and helping teachers understand the reasons and need for institutional change more explicitly. With regard to the potential of action research to extend its impact beyond the classroom walls, Burns speculated: "Can action research go beyond the individual classroom to have a broader impact on the institution in which it takes place? In my view it can when it occurs in a collective way" (p. 14). However, Burns does not back up this personal opinion with any substantive data.

According to Sagor (1991a), school-based action research has the potential not only of enhancing the professional lives of teachers but also of impacting student achievement. Sagor described an action research project in which a single teacher implemented new teaching strategies in her classroom and subsequently watched test scores for low achievers rise above those of their high achievement counterparts. This teacher's results motivated others in the school to adopt similar strategies. However, Sagor did not present evidence as to whether the implementation of these practices produced similar results school-wide.

Using survey and interview techniques, Sagor (1991b) also assessed the impact of action research in 5 schools on teachers' perceptions of school culture and student achievement. Sagor concluded that, after two years of involvement in action research, all teachers agreed that their schools were getting somewhat better. Collegiality had become more meaningful and satisfying, teacher self-esteem and self-efficacy had increased, and teachers felt that the workplace was less isolating and more fulfilling. Data on student

levels of academic achievement, social behavior, and attendance was less encouraging, causing Sagor to surmise that the action research intervention was too new at these schools to be responsible for changes in achievement patterns. However, Sagor maintained that “collaborative action research has potential to foster school effectiveness” (p. 59).

Research has suggested that the action research model improves school culture and teacher attitude. However, there is very little evidence at this point regarding the effects of action research on sustained changes in teaching or student performance. Even evidence of the impact of action research on teachers’ attitudes and school climate tends to be anecdotal. Despite this fact, action research is the “buzzword” of the staff development field at this time. However, it must be kept in mind that enthusiasm cannot compensate for a lack of solid evidence. At present, data regarding the true impact of action research is lacking.

Organizational development (OD) model

The organization-development model involves both teachers and other school staff. Deriving from the belief that many problems are a function of flaws in a school’s organizational structure, this model focuses on everyone in a school. Ideally, the end result of an organizational development staff development model is the resolution of the initial problem, as well as increased capacity on the part of all school staff to solve future problems and engage in a process of continuous improvement (Gall & Vojtek, 1994).

In the OD model, the staff developer is usually seen as a “consultant” and a “facilitator.” As a consultant, the key responsibility of the staff developer here is

facilitating organizational change and development by working with individuals who comprise the organization. Other responsibilities include sharing skills and knowledge to build the capacity of others and help teachers be successful in their work by coaching, advising, considering consequences, and suggesting alternatives (Killion & Harrison, 1997, p. 8). According to Killion & Harrison (1997), a facilitator is one who “makes things happen with ease” (p.10). Therefore, staff developers in OD projects are charged with facilitating both tasks and processes. A “task facilitator” orchestrates a project or assists a group in completing a task. A “process facilitator” focuses on the interactions among group members.

Research on the Organizational Development Model

Few evaluations of organizational development (OD) initiatives in education have been undertaken. Those that do exist do not generally report student-level results or changes in teaching. Rather, they tend to report changes in school climate and participant attitudes. One example of this is Runkel and Schmuck’s (1985) analysis of a six-year OD project in one school district. Designed to impact administrators’ working relationships and problem-solving processes and staff norms and processes, this project achieved an uneven level of success. Via a series of questionnaires, Runkel and Schmuck concluded that:

- Schools participating in OD scored higher on tests of communication than did non-participating schools.

- Staff that established collaborative problem-solving and decision-making structures showed heightened communication skills for a year than did non-participating schools. After 2 years, however, this effect declined.
- A small amount of training in communication and organizational skills did more harm than good. While schools that received 16 or more hours of communication training and 24 or more hours in organizational skills training showed increases in these skills, schools that received less than these amounts declined in communication and organizational skills.
- Staff participating in OD were more likely to communicate with others when choosing teaching methods.
- Elementary schools with OD training and consultation used team teaching more successfully than did those schools without training, suggesting that OD efforts may increase a school's readiness for innovation and its ability to make the changes needed to permit innovations to endure (Fielding & Schalock, 1985).

On the other hand, Neuman, Edwards, and Raju (1989) meta-analyzed 126 organizational development studies conducted in management and the behavioral sciences. Using job satisfaction and attitude change as independent variables, Neuman, et al. concluded that organizational development initiatives affect attitudes more than satisfaction. Furthermore, team building (interventions in which intact work groups learn to increase their skills for effective teamwork), followed closely by lab training (4-5 day workshops designed to identify group problems, identify solutions, and plan implementation procedures) were found to be the most effective organizational

development techniques for modifying satisfaction and other attitudes. These researchers assert that team building and lab training encounter a high degree of success because they “provide employees with the most direct experiences of self-development and improvement” (p. 481). It must be pointed out that Neuman, et al. caution readers of this study about generalizing these findings due to the non-random selection of techniques, variable rigor of the studies analyzed, and the study’s frequent failure to account for much of the variability of the effect sizes.

In summary, research has not traditionally linked OD efforts with sustained changes in teacher performance or student achievement. In addition, research suggests that a little bit of OD is worse than none at all. While OD does appear to improve participants’ communication skills and attitudes, the “fuzzy” nature of this model does not facilitate the identification of certain OD techniques that are more effective than others. Furthermore, since this model is so closely linked with the facilitator/consultant, it seems likely that the success of this approach depends heavily on the skills of the staff developer.

Development/improvement process model

The development/improvement process model (Sparks & Loucks-Horseley, 1990) is also known as the change process model (Gall & Vojtek, 1994). This model is typically used when teachers are required to develop or adapt curriculum, design programs, or engage in systematic school improvement processes that have as their goal the improvement of classroom instruction and/or curriculum. Successful completion of these projects typically requires that teachers acquire specific knowledge of skills or

develop a product that results in significant student learning. Staff development activities under this model may include reading relevant literature, group discussion, observation, training, trial and error, experiential learning, organizational development, and more. As such, staff developers in this model serve multiple roles. They can be trainers, resource providers, coaches, facilitators, consultants, depending on the needs of the school and staff members and the activities undertaken. In addition, Killion and Harrison (1997) suggest that the staff developer should be seen as a “catalyst for change” who suggests new ideas and questions existing practices. He or she must promote and guide continuous analysis and reflection among members of the organization and initiate alternatives to current practices (p. 14).

According to Gall and Vojtek (1994) the development/improvement process model is the most complex of all staff development models. They also stipulate that it takes the longest time to be achieved; for example, Gall and Vojtek cite research (Fullan and Steigelbauer, 1991) indicating that moderately complex changes that can require from three to five years, while major changes can require from 5 to 10 years. This model is further complicated by the fact that several of the staff development models presented above are included in various stages of the change-process model.

One assumption of this model is that adults learn most effectively when they have a need to know or a problem to solve. Another assumption is that teachers best understand what is required to improve their own performance. A final assumption is that teachers acquire important knowledge and skills through their involvement in school improvement or curriculum development processes. It is believed that involvement in

these processes is likely to cause changes in attitudes or acquisition of skills. In this model, most staff development activities are conducted with groups of teachers, increasing the probability that teachers will become more cohesive and share ideas about teaching and learning in general, as well as those focused on the development task of hand.

According to Gall and Vojtek (1994), the development/improvement process model is appropriate when desired improvements require systemic change at the school district level. Gall and Vojtek (1994) cite staffing or curriculum changes as typical school improvements necessitating system change. This model has three stages: initiation, implementation, and institutionalization.

In the initiation stage, a decision about whether to adopt the proposed innovation is made. According to Gall and Vojtek (1994), teachers need to receive continuous, personalized information about the innovation. In the implementation stage, the innovation is put into action in a particular school or other setting. Effective staff development in this stage involves a combination of concrete, teacher specific training activities, ongoing continuous assistance and support during the process of implementation, and regular meetings with peers and others.

Proponents of the development/improvement process model find the Concerns-Based Adoption Model (Hall & Loucks, 1978) of change useful both in understanding the change process and implementing appropriate staff development activities at various points in the process. In fact, Hall and Loucks (1978) developed an instrument, the Stages Of Concern Questionnaire, that is often used by staff developers in a change

process to identify the type of concern that each teacher is experiencing at a particular point in time in the change process and design appropriate staff development activities to address teacher concerns.

During the final stage of the development/improvement process model, a decision is made to continue using the systemic innovation indefinitely. According to Gall and Vojtek (1994), staff development is necessary to ensure that the innovation continues to be used as intended.

Research on the Development/Improvement Process Model

The literature on the development/improvement process staff development model is positive, in terms not only of changing teacher attitudes and beliefs, but also of impacting student achievement. For example, Sparks, Nowakowski, Hall, Alec, and Imrick (1985) reported the success of the Staff Development for School Improvement project, a 3-year, teacher-directed school improvement project that was implemented in 19 schools in Michigan. Using training, organizational development, and individually guided staff development techniques, teachers in these schools collaborated to participate in activities such as: examining the state assessment objectives, analyzing curricula, introducing and practicing effective teaching techniques, revising school mission statements, and participating in self-selected professional growth activities. At the two schools cited in Sparks et al.'s research report, above-average student achievement on the state reading test improved from 77.6% to 97.5% and from 72% to 100% in 2 years. Because staff at the schools remained stable and no other programs were initiated during this time period, Sparks, et al. (1985) attributed these remarkable gains to the staff

development program and not outside factors. Unfortunately, Sparks, et al. failed to take into account that these gains could possibly have been attributed to changes in the student population, improved student familiarity with the format of the state test, increased teaching to the test on the part of teachers, or a number of other factors.

In a study of the common characteristics of effective school programs, Miles, Farrar, and Neufeld (1983) conducted telephone interviews with program directors and collected program-relevant materials and evaluation reports from 39 schools that had successfully implemented new programs. Interestingly, the staff development activities these schools had in common consisted of strategies that are commonly used in the development/improvement staff development model: data collection and feedback, school-wide planning teams and task forces, shared decision making, consulting/technical assistance, and intensive training. Unfortunately, data from and about teachers is lacking in this study, rendering it impossible to identify the impact of the above staff development strategies on teacher behavior.

In their 3-year study of a collaborative curriculum project involving 750 teachers, Kimpston and Rogers (1987) obtained pre- and post- measures of participating and control teachers' behaviors and attitudes toward curriculum plans and planning. Results revealed that "a few strong and significant changes appeared to occur in the participants' general attitudes and behaviors toward curriculum plans and planning over the three-year period" (p. 14). While these results are positive, Kimpston and Rogers (1987) caution readers to keep in mind the difficulty of generalizing case study data and the potential for inaccuracy when collecting data based on self-report.

According to Aschbacher (1993), Borko (1997), and Shepard (1997), the implementation of regular teacher discussion groups *plus expert consultation* is effective at helping teachers cope with the demands imposed on them by new curriculum frameworks and national standards in math and literacy. In an attempt to empower teachers to select and develop performance assessments congruent with school-wide assessment reform, researchers met weekly with teachers who were trying out new assessment techniques. These weekly sessions allowed teachers to talk with colleagues about the logistics and underlying rationale of these new strategies. An assessment expert was present at each session to offer teachers support in learning the conceptual basis behind the intended reforms and “supply good examples in response to teacher-identified topics” (p. 11). After one year of such sessions, Shepard (1997) reported that teachers were using more activities closely aligned with standards, students had a greater understanding of why teachers graded the way they did, and teachers felt they had greater knowledge about what students could do. In addition, students showed appreciable gains in math achievement, as compared to students in matched-control schools. Caution must be taken in interpreting the results of these studies, however, due to extremely small sample size and the fact that the researchers themselves (Aschbacher, Borko, and Shepard) were the “expert consultants” described in the studies.

Borko and Shepard’s recommendation regarding expert assistance echoes findings by Crandall and Loucks (1983), who conducted nearly 500 interviews and surveyed 4,000 teachers, school staff, and staff developers involved in federal school improvement efforts. Specifically, Crandall and Loucks asserted that “new practices entailing a

significant amount of change live or die by the amount of personal assistance they receive” (p. 11). According to these researchers, effective, sustained assistance, which may be internal or external, help teachers master new practices, in turn leading to student gains and the achievement of school improvement goals. In fact, Crandall and Loucks postulate that external consultants (e.g., university researchers) combined with internal assistance (e.g., peer assistance) dramatically increases the amount of change that is brought about by either source alone.

In his study of the staff development and instructional improvement efforts of Community District 2, New York City, Elmore (1996) concluded that a “loosely-connected, constantly-evolving set of [staff development] activities held together by a common theme of instructional improvement” (p. 24) was key to changing teaching practice. As such, he recommended a set of multi-year, specific staff development activities implemented by District 2 and demonstrated (by anecdotal evidence) as having extensive effects on teaching practice. The first of these activities includes the implementation of a “professional development laboratory” whereby experienced “Resident Teachers” agree to accept a certain number of teachers as visitors in their classroom for 3 weeks of intensive observation and supervised practice. At the same time, experienced and highly qualified “Adjunct Teachers” substitute in the visiting teachers’ classrooms and spends a considerable amount of time observing and practicing in Resident Teachers’ classrooms. Adjunct and Resident teachers subsequently support the development of new practices by observing visiting teachers’ classrooms and consulting on issues of practice.

Other staff development activities implemented in District 2 and recommended by Elmore (1996) include the provision of intensive, long term instructional consulting services to teachers. Under this model, consultants work one-on-one with 8-10 teachers for blocks of three to four months each and with grade-level teams and larger groups of teachers during planning time, after lunch, and at school. This model also includes substantial inter-school visitation, peer networks and off-site training combined with support or direct assistance in the classroom on the instructional changes focused on in training. Finally, frequent formal and informal visits by the superintendent and district personnel, involving conversations with principals about specific issues in the school, a walk-through and visits to several classrooms, debriefing sessions with principals, and a follow-up letter summarizing the results of the visit and the agreed-upon actions are recommended as a way of signifying that the intended instructional improvement is supported at the district level. While stating that anecdotal evidence was available to support the effectiveness of such a staff development strategy, Elmore did conclude that more information on changes in teaching practice and their relationships to student learning were necessary to justify such a high cost, labor intensive approach.

Based on the above research findings, the key to the development/improvement model seems to be time and technical assistance. The research cited suggests that successful development/improvement takes years and is greatly assisted by sustained expert consultation. It is not clear, however, whether certain staff development strategies within this model lead to more positive results than others, rendering it difficult to use these research findings to design future development/improvement efforts.

Professional Development School model

Darling-Hammond (1998) likens Professional Development Schools (PDSs) to teaching hospitals, in that “they provide sites for state-of-the-art practice that are organized to support the training of new professionals, extend the professional development of veteran teachers, and sponsor collaborative research and inquiry” (p. 9). Alternatively known as professional development schools, professional practice schools, clinical schools, laboratory schools, school-college partnerships, and induction schools, PDSs are collaborations between schools and universities where opportunities are provided for novice teachers to learn from expert practitioners, while experienced teachers act as mentors, university adjuncts, school restructurers, and teacher leaders (Darling-Hammond, 1995).

PDSs are designed to be outstanding public schools, cooperatively established and maintained by schools of education and selected school districts. They are also intended to strengthen knowledge and practice in schools by providing exemplary sites for research, experimentation, inquiry, evaluation, and eventual dissemination of innovative programs and effective practices. Finally, PDSs are meant to serve as models of promising and productive structural relationships among school staff (Abdal-Haqq, 2000, pp. 1-3).

According to Darling-Hammond (1995), PDSs provide valuable staff development opportunities for veteran teachers, in that they learn more about the theory and practice of teaching as they become teacher educators. At these sites, curriculum development is always in development, adaptive teaching strategies are regularly

discussed, shared decision making is the norm, and leadership stems from teachers' interests and expertise, rather than formal roles and titles. Consequently, PDSs are thought to have the potential to "transform the entire educational enterprise by changing teaching, schooling, and teacher education simultaneously" (Darling-Hammond, 1995, p. 87).

Research on the Professional Development School model

PDSs have existed in the United States since 1887, when the Horace Mann School at Teachers College, Columbia University, was established. PDSs reached their peak in 1964, with 212 schools. However, PDSs were in decline by 1969, as research results began to surface indicating that PDSs had become too far removed from the main stream of school life to be credible, school-university collaboration was more talked about than implemented, and the effects of PDSs on teachers and students was unexamined (Stallings & Kowalski, 1990).

In the 1980s interest in professional development schools was rekindled, with plummeting test scores in the 1970s and the publication of A Nation at Risk in 1983. In addition, reports published by the Carnegie Forum on Education and the Economy (1986) and the Holmes Group (1986) recommended the establishment of clinical schools to prepare teachers to meet professional standards on subject-matter knowledge, pedagogical understandings, and application of knowledge in the classroom. The goal of these new PDSs was to link school and college faculties to provide the best possible learning environment for teacher preparation and professional development (Stallings & Kowalski, 1990).

As of 1996, there were approximately 344 documented PDSs in the United States, representing 84 partnerships between schools and universities. Seventy-one percent of these PDSs were relatively new, having been initiated between 1991 and 1996 (Professional Development Schools at a Glance, 2000).

At present, there is some evidence from descriptive and qualitative studies, as well as anecdotal evidence, that PDSs are effective at increasing teacher satisfaction and efficacy, as well as changing their attitudes and beliefs about teaching. However, this evidence is scant. For example, Clinard, Ariav, Beeson, Minor, and Dwyer's 1995 study of the impact of participating in a PDS on cooperating teachers' teaching and professional lives revealed new attitudes about students and teaching, increased teacher reflection, improved confidence in dealing with peers, better social interaction and communication skills, and increased pride in the teaching profession. These results were based on descriptive observations of large group interaction and written feedback from 172 cooperative teachers in California PDS. Other studies have indicated that PDSs are effective at increasing teacher morale, retention, and expectations of students, as well as improving school-university relations and the background knowledge of teachers (Stallings & Kowalski, 1990).

At present, there is little evidence that PDSs ultimately alter the practices of experienced teachers or enhance student learning. In addition, some researchers (Teitel & Abdal-Haqq, 2000) maintain that PDS implementers and advocates are reluctant to evaluate the impact of PDSs. As a result, systematic documentation and assessment is infrequently built into program design (Professional Development Schools: Frequently

Asked Questions, 2000). Consequently, the evidence behind PDSs has not reached a point where substantive conclusions regarding the impact of this staff development model (on classroom practice and student learning) can be made.

Organizational context model

The final staff development “model” posited by Sparks and Loucks-Horsely (1990) is, in fact, hardly a model at all. Rather, Sparks and Loucks-Horsely’s “organizational context model” identifies organizational context features that increase the probability that a staff development program will encounter success. Features of this “model” describe organizational characteristics such as: norms of collegiality and experimentation (Little, 1982), clear goals and a stable organizational structure (Joyce & Showers, 1983), active support by principals and district administrators (McLaughlin & Marsh, 1978), principal involvement in staff development (Stallings & Mohlman, 1981), strong pressure for implementation (Huberman, 1983), avoidance of “innovation overload” (Anderson & Odden, 1986), involvement of participants in key decisions about staff development (Lieberman & Miller, 1986), and clear, common goals (Ward & Tikunoff, 1981; Anderson & Odden, 1986). Furthermore, teacher characteristics believed by some researchers to impact school change are generally included in this model. These teacher characteristics typically include number of years of teaching experience and teachers’ sense of self-efficacy, among others (Berman & McLaughlin, 1977, 1978; Sparks, 1983; Gliessman, Pugh, Dowden & Hutchins, 1988; Kimpston & Rogers, 1987).

Research on the Organizational Context Model

Research on the organizational and personal factors that may influence the implementation of innovations is extensive. In 1976, Berman and McLaughlin, authors of “the Rand Study,” wrote that “project outcomes depended more on the characteristics of the project’s setting than on any other factor” (p. 361). The 4-year, 2-phase Rand study, which examined approximately 300 educational innovations through surveys of 852 administrators and 689 teachers and multiple observations of programs in operation at two different time points, concluded that the quality of the school’s organizational climate played an important role in project implementation and continuation (Berman & McLaughlin, 1977, 1978). In addition, this extensive study revealed that the quality of working relationships and the number of project meetings at the school level affected the number of goals achieved and whether or not staff development projects lasted.

Based on an extensive review of literature on innovative and non-innovative schools, as well as 45-minute interviews with ten practicing school-based change agents, Manning (1976) identified a series of innovative (and non-innovative) school characteristics which, when measured, could be used to predict and diagnose a school’s readiness for change. In particular, characteristics associated with innovativeness were related to school-based staff, patterns of communication, organizational climate, and school-community relations. For instance, school-based staff in innovative schools were found to have a high degree of interaction and peer support. In addition, staff members tended to experiment with new practices and support innovation. Similarly, communication between teachers and principals in such schools was direct, consistent,

honest, and reciprocal. These schools were cooperative, provided teachers with adequate resources and planning time, and were not characterized by staff divisiveness. Finally, Manning asserted that communities in which innovation was successful were characterized by widespread parent and community involvement and support.

Little (1981) conducted a much cited study in six desegregated schools selected to represent a wide range of involvement in staff development initiatives. Designed to explore the ways in which the social organization of schools as workplaces bear on teachers' involvement in formal and informal learning "on the job," this focused ethnography concluded that continuous professional development is most thoroughly achieved when teachers engage in consistent collegial interactions characterized by: continuous discussion of classroom practice, mutual observation and critique, shared efforts to design and prepare curriculum, and shared participation in instructional improvement.

Little (1981) also identified specific school and teacher characteristics that were found to be typical of schools that were not only heavily involved in staff development but also successful at school improvement efforts. For instance, she found that the sheer frequency of interaction among teachers was an indicator of high involvement and high success. Furthermore, interactions about teaching were seen as reciprocal, even when they involved people of different status (principal versus teacher) or different roles (teacher versus staff development consultant). Finally, successful, involved schools tended to be inclusive; most of the faculty participated in innovation. In addition, Little contended that staff development had the greatest prospects for influence "where

analysis, evaluation, and experimentation are treated as tools of the profession, designed to make work better (and easier), and where such work is properly the work of the teacher” (p. 339).

According to Elmore (1996), the creation of a culture of shared values around instructional improvement will bind the work of teachers and administrators into a coherent set of actions and programs. Drawing on interviews from individuals involved in a staff development and instructional improvement project in New York City’s Community District 2, Elmore identified a set of organizing principles about the process of systemic change and the role of professional development in the change process. Central to these principles are the issues of shared expertise as the driver of instructional change; collegiality, caring, and respect among teachers and principals; and getting people together to share and generate ideas. Elmore asserted that there was ample “anecdotal evidence” of the effects of this culture of shared values on teachers in Community District 2, while at the same time conceding that both the proportion of teachers who have changed their practice, as well as the relationship of changes in teaching practice to student learning, were unknown.

The Rand Study (Berman & McLaughlin, 1978) found that the “the support of the building principal for a special project was directly related to the likelihood that staff would continue to use project methods and materials after special funding is withdrawn. Furthermore, principal support positively affected project implementation” (p. 81). Likewise, Little (1981) concluded that teachers in successful, innovative schools tended to view the principal as “an active endorser and participant in collegial work” (p. 337).

These findings were supported by Gall, Fielding, Schalock, Charters, and Wilczynski (1984), who experimented with 49 teachers in 15 schools on the teaching of math. Teachers were non-randomly assigned to a principal-involvement group, a regular in-service group, and a control group. Principals in the principal-involvement group attended training sessions, observed lessons, and met with participating teachers about the staff development project. This group of teachers showed greater ($p=.07$) implementation than both the in-service and control groups, suggesting that principal involvement played a key role in the transfer of new skills to the classroom.

Based on data gathered from 4,000 questionnaires and approximately 500 face-to-face interviews with teachers, school staff, and staff developers involved in federal school improvement efforts, Crandall and Loucks (1983) concluded that “forceful leadership is the factor that contributes most directly and surely to major, effective changes in classroom practice that become firmly incorporated into everyday routines” (p. 10). This necessitates that principals be competent, forceful, and knowledgeable about new practices. Similarly, Sagor (1991b) concluded that collaborative action research could not overcome the handicaps of inadequate principal leadership.

Drawing on the experience of Community District 2, New York City, Elmore (1997) identified the principal as the key actor in instructional improvement, stating that the ideal principal is highly engaged in and focused on instructional practice and actively engaged in the creation of a professional community in a school. In addition, successful principal-school combinations and the care that went into pairing them led Elmore to

conclude that potential principals must be carefully matched with the unique set of characteristics each school presents.

Much speculation has been made about whether years of teacher experience has an effect on the success of staff development efforts. The Rand Study revealed a consistent negative relationship between teachers' years of experience and project goals achieved, teacher change, and student improvement (Berman & McLaughlin, 1977). In fact, Berman and McLaughlin (1978) suggest that teachers "peak out" after five to seven years of teaching—either maintaining their level of effectiveness or actually becoming ineffective. In an attempt to identify the effects of variables that cut across staff development methods, Gliessman, Pugh, Dowden, and Hutchins (1988) reviewed 26 experimental studies to identify the effect of trainee characteristics on skill changes through training. Among their conclusions was the discovery that the mean effect size (.91) for inexperienced teachers was not significantly different ($z=1.43$, $p>.05$) from the mean effect size (.65) for experienced teachers. On the other hand, Kimpston and Rogers (1987) found that the more experience teachers had with collaborative curriculum planning (an incarnation of the development/improvement process model), the more likely they were to alter their attitudes toward curricular topics following a collaborative curriculum planning initiative.

In the Rand Study (Berman & McLaughlin, 1977), teachers' sense of self-efficacy emerged as a powerful predictor of project goals achieved, improved student performance, teacher change, and continuation of project methods and materials. Interestingly, teacher self-efficacy was not significantly related to years of experience

(Berman & McLaughlin, 1978). Interviews and case studies conducted in Sparks' 1983 study of in-service training, teacher characteristics, and teacher change also supported the possibility that teachers' self-efficacy might predict implementation and change.

In summary, research has produced very consistent results with regards to the organizational context model. In fact, there is substantial consensus on the organizational and personal factors that impact school-wide and teacher change. This model is useful largely in predicting an institution's or a teacher's receptivity to innovation and the likelihood that an institution or a person will change. However, the model does not consist of a set of strategies or activities that will bring about change. As such, the organizational context model can only be used in conjunction with another model, one that is based on concrete steps designed to bring a person or school from point A to point B.

Gaps in the Staff Development Research

According to Sparks and Loucks-Horsley, the purpose of staff development is to "improve student learning through enhanced teacher performance" (p. 235). As such, the ultimate measure of the success of a staff development initiative is the degree of impact the initiative has on both classroom teaching practices and student learning.

Furthermore, staff development is viewed as the essential component of school change initiatives in the current era of educational reform. Millions of dollars are spent on staff development annually, and those responsible for selecting staff development programs are faced with a "giant academic bazaar" (Stout, 1996) in which a multitude of staff development strategies, materials, and timeframes are purported to address any single

staff development objective. From the point of view of increasing student learning, efficiency and cost-effectiveness, it would be helpful for school personnel to be able to identify the particular staff development technique(s) that are most likely to impact teacher performance and student achievement in their unique setting .

A review of the research on staff development reveals that the success of staff development has typically been assessed according to its impact on at least eight levels: teacher reaction, changes in attitude, increased awareness, acquisition of concepts and organized knowledge, demonstration of principles and skills, classroom transfer, impact on students, and organizational change.

Historically, staff development has been deemed “effective” if teachers respond positively to it. As such, this *reaction* level of impact has been commonly assessed through the use of “happiness quotients” or “smile sheets” completed by teachers at the end of a training session. While some criticize this level of impact as merely an indication of “the entertainment value of an activity” (Guskey, 1998, p. 5), others insist that positive reactions to staff development are a necessary prerequisite to higher levels of impact (Kirkpatrick, 1959, 1977).

Similarly, many staff development activities are considered effective if they result in a change in attitude on the part of the teachers who participate in them. As such, a common level of impact that is measured in staff development studies can be labeled *attitude*. The main idea behind the quest for teacher attitude change is the assumption that changes in attitude will lead to changes in practice and, ultimately, to increases in

student performance. Unfortunately, however, the connection between attitude change and subsequent student achievement is assumed and rarely dealt with directly.

In 1980, Joyce and Showers (1980) presented what was to become a widely-accepted four-level classification of levels of impact of any staff development initiative. Joyce and Showers label these levels: *awareness*, *concepts and organized knowledge*, *principles and skills*, and *transfer*. When a staff development initiative helps participants become aware of the importance of a teaching strategy or approach, it achieves an impact at the awareness level. On the contrary, staff development that provides teachers with relevant concepts and knowledge related to a change impacts them at the level of concepts and organized knowledge. When teachers gain the principles and skills necessary to act in a new way, they have been impacted at the principles and skills level.

At Joyce and Showers' final impact level, transfer, the concepts, principles, and skills of a new mode of action are transferred to individual teachers' classrooms. At this point, teachers use new strategies in their classrooms and integrate them into their existing repertoires. Researchers such as Joyce and Showers (1980) are clear that changes in student achievement cannot be expected until a staff development initiative has impacted teachers at the transfer level. As such, an additional *student-level* of impact is warranted and is often thought to be the most important indicator of an initiative's effectiveness. After all, the primary responsibility of teachers is to educate students, and staff development is designed to help teachers perform their jobs better. However, using student learning measures as the principal criterion is exceptionally rare when

determining effective professional development (Guskey & Sparks, 1991; Sparks, 1995; Guskey, 1997).

Additionally, staff development has been deemed effective or successful if teacher/administrator interactions or school climate is altered in a positive way; this level of impact can therefore be called *organizational*.

A review of the literature related to the impact of staff development reveals that the objectives of most staff development efforts have focused on teachers' acquisition of or refinement of skills. On the other hand, the majority of staff development evaluations have focused on teacher reactions and attitudes. Obviously, there has been a consistent mismatch between staff development goals and the questions asked as part of evaluation. More important outcomes, such as the transfer of training to the classroom or changes in student achievement, are infrequently assessed as a measure of the success of various staff development models and strategies.

Further complicating this situation is the fact that staff development models are frequently combined in professional development initiatives. The OD and development/improvement process models are particularly plagued by this reality. For instance, a single OD project might include training, coaching, discussion groups, and other diverse activities. However, most studies of staff development have not been designed to measure the level(s) of impact of each model used within a larger, more inclusive model, rendering it difficult to determine which staff development strategies or combinations of strategies are responsible for programmatic successes and failures.

In addition, certain staff development models are subsumed in others. For example, participants in coaching and action research projects must be trained to participate these efforts prior to carrying them out. The success of the coaching or action research, therefore, could very well be linked with the quality of initial training. However, this is a point that is not reported or discussed in most studies of coaching or action research.

Finally, individual studies and meta-analyses tend to produce inconsistent results. According to Guskey (1995), this can be explained by the fact that meta-analyses synthesize evidence across studies by decontextualizing data. The elimination of contextual information related to the uniqueness of individual settings, therefore, “waters down” the data and leads to few significant, conclusive results.

Due to the nature of the staff development research conducted to date, Guskey (1997) states that:

“...we’re still not sure precisely which elements contribute most to effective professional development, what formats or specific practices are most efficacious, or precisely how professional development contributes to improved teaching and learning” (p. 1).

Historically, the focus of staff development research has been on identifying the “one right answer” or optimal combination of staff development model(s) that most changes teachers attitudes or improves their skills. Unfortunately, this focus has not yielded consistent, usable results. What is needed now is a study of the conditions—the staff development formats contextual characteristics—that are likely to have a positive effect on classroom practice and student learning in a variety of settings. While a single,

generalizable answer may not be yielded, the results of such a study should provide staff developers and school staff with knowledge that can be used to design individualized staff development programs with the largest possible likelihood of success

Summary

This chapter has reported many studies and reviews of the literature pertaining to the nature of change, the history of staff development for teachers in the US, staff development models, and research on effective staff development strategies. These studies have formed the immediate background and context for the design of the current investigation. This design is described in Chapter 3.

CHAPTER 3

Research Design

Overview

This chapter discusses the design of this study and includes a description of research hypotheses, teacher and student populations and samples, survey instruments, teacher interview procedures, independent and dependent variables, methods used for collection and treatment of data, and limitations of the present study.

Research Hypotheses

This investigation attempted to address the following: Which staff development strategies and contexts have a positive impact on teachers' attitudes and practices? For the purposes of this study, *staff development* was defined as "any teacher activity, conducted individually or in a group, for the purposes of altering the professional practices, beliefs, and understanding of teachers and ultimately bringing about an improvement in student learning" (based on a definition provided by Guskey, 1986, p. 5). As such, staff development could take on a variety of forms, such as workshop attendance, peer coaching, modeling, independent reading, study groups, journal writing, teacher discussion, action research, project meetings, and more.

To address this question, the study focused on a sample of 89 Rhode Island elementary teachers in 11 schools across 5 districts, all of whom were participating in a state-sponsored staff development project designed to assist them in acquiring the knowledge and skills necessary to develop and implement standards-based math instruction and assessment in their classrooms.

The investigation tested the following null hypotheses:

Hypothesis 1: Changes in teachers' level of concern toward standards-based instruction and assessment in math cannot be predicted by

- a) teacher role in the staff development project (i.e., lead vs. participating teacher)
- b) size of the schools in which teachers work
- c) total number of years teaching experience
- d) teachers' sense of self-efficacy
- e) quality of school-wide communication
- f) degree of principal directiveness
- g) degree of principal supportiveness
- h) quality of school/community relations
- i) quality of organizational climate
- j) staff development activities in which teachers engage

Hypothesis 2: Changes in teachers' intensity of concern toward standards-based instruction and assessment in math cannot be predicted by

- a) teacher role in the staff development project (i.e., lead vs. participating teacher)
- b) size of the schools in which teachers work

- c) total number of years teaching experience
- d) teachers' sense of self-efficacy
- e) quality of school-wide communication
- f) degree of principal directiveness
- g) degree of principal supportiveness
- h) quality of school/community relations
- i) quality of organizational climate
- j) staff development activities in which teachers engage

Hypothesis 3: Changes in teachers' reported use of standards-based instruction and assessment in math cannot be predicted by

- a) teacher role in the staff development project (i.e., lead vs. participating teacher)
- b) size of the schools in which teachers work
- c) total number of years teaching experience
- d) teachers' sense of self-efficacy
- e) quality of school-wide communication
- f) degree of principal directiveness
- g) degree of principal supportiveness
- h) quality of school/community relations
- i) quality of organizational climate

- j) staff development activities in which teachers engage

Independent Variables

The independent variables in this study were: a) teacher role in the staff development project (i.e., lead vs. participating teacher); b) size of the schools in which teachers work; c) total number of years teaching experience; d) teachers' sense of self-efficacy; e) quality of school-wide communication; f) degree of principal directiveness; g) degree of principal supportiveness; h) quality of school/community relations; i) quality of organizational climate; and j) types of staff development activities engaged in. Each of these variables is described below.

The variable corresponding to teacher role in the project was determined by whether teachers had elected to become lead or participating teachers. Apart from possible innate factors such as personal motivation, openness to innovation, and leadership qualities, the primary difference between lead and participating teachers was that lead teachers from each school attended three 1-1/2 day workshops with lead teachers from all of the other schools in the staff development initiative. Participating teachers did not attend any such workshops.

The variable, *school size*, referred simply to the total number of students enrolled in schools in which teachers worked. Similarly, the variable corresponding to total number of years teaching experience referred to the number of years teachers had taught in any school, as reported by teachers themselves.

Teachers' sense of self-efficacy referred to the general belief in one's ability to perform a variety of tasks (particularly in the face of adversity) as well as one's willingness to initiate and persist in the completion of a task.

The variable, *quality of school-wide communication*, was characterized by the degree to which direct, face-to-face communication occurred among teachers, principals, consultants, and administrative staff via individual and group meetings. It also included the extent to which principals, lead teachers, and school consultants sought teachers' opinions.

Two variables focused directly on the principal in each teacher's school. The first variable, the *degree of principal directiveness*, denoted the degree to which principals single-handedly made all class and work scheduling decisions, as well as the control principals exercised over agendas and activities conducted at faculty meetings. In addition, the principal directiveness variable included how much principals tended to correct teachers' mistakes and allow their rules to be questioned.

In contrast, the *principal supportiveness* variable referred to the willingness of principals to work with teachers during and outside the school day, as well as the degree to which s/he shared his/her educational knowledge with teachers. Furthermore, this variable included the clarity of communication from the principal to teachers, principal assistance in settling teacher differences, and the degree to which principals offered teachers helpful, constructive criticism.

The *quality of school/community relations* can be described as the awareness and involvement of parents and other community members in standards-based reform and

innovation in general. This variable also referred to whether or not parents felt comfortable offering their opinions to the school administration and the extent to which the school system was sensitive to community opinions. Finally, this variable included the degree to which parents were informed of and attended school events.

The *organizational climate* variable dealt with the degree to which isolated subsystems of teachers exist within a school. In addition, this variable encompassed the organization's openness to innovation and the strength of channels and procedures for dissemination.

Several variables were used to describe the *types of staff development activities* in which teachers were engaged. These variables corresponded to activities listed by teachers in their professional development activity log. Types of staff development in which teachers engaged included *action research*, *workshops*, *meetings*, work with educational *consultants*, carrying out *standards-based activities* in the classroom, *independent research*, *professional reading*, *meetings/discussions* with colleagues, *curriculum development*, and examining *student work*.

Action research refers to “the process by which teachers systematically examine their own practice using research techniques” (Watts, 1985). As such, the *action research* variable was distinguished by the following essential activities: identifying “essential questions,” or research topics, collecting data to address essential questions, analyzing data, and compiling final reports or projects to showcase one's action research results.

The *workshop* variable was characterized by participation in school-level in-service sessions, district, state, or national staff development workshops, local and/or national education conferences, graduate level education courses. This variable also included the viewing of training videos.

The *meetings/discussions* variable pertains to all formal and informal meetings and/or discussions with one or more teachers, supervisors, or school officials on the subjects of standards-based instruction and assessment and/or school improvement. This variable also includes preparation for meetings.

The *consultants* variable corresponds to any group or individual activity undertaken with the assistance of an expert consultant assigned (through the staff development initiative) to work with teachers at the school site.

The *standards-based activities* variable corresponds to the implementation of standards-based instruction and assessment in the classroom. For most teachers, this involved experimentation with new types of materials and tasks, as well as implementing their school's new standards-based math curriculum for the first time. Preparation for the implementation of standards-based activities was integrated in this variable as well.

The *independent research* pertains to teachers' efforts to identify standards-based resources, such as instructional tasks, reading materials, and information about other schools' experience implementing standards-based instruction and assessment in mathematics. The majority of teachers' independent research consisted of research on the Internet and the identification of useful web sites.

The *professional reading* variable consisted of reading academic literature related to standards and the teaching of mathematics, as well as reading standards documents themselves (i.e., the New Standards) and standards-based curriculum books (for purposes other than planning lessons).

The *curriculum development* variable includes the following activities: developing/revising standards-based tasks and assessments, designing rubrics, and aligning curricula to standards.

The *student work* variable corresponds to ways in which teachers examined student work/performance, outside of the traditional activities of grading papers or tests. Conducted primarily in small groups, these activities incorporated under this variable included saving student work samples and comparing student performance over time; deliberately observing students in the process of completing math tasks and taking note of relevant details surrounding their performance; videotaping students in class and scrutinizing the strategies individual students employed; scoring standards-based assessments using rubrics; and setting benchmark papers. All of these activities could take place individually or with groups of teachers.

For each staff development activity engaged in, the number of hours each teacher devoted to that activity was utilized in the analysis of the data.

Dependent Variables

Three dependent variables were used in this study: 1) change in level of concern about standards-based instruction and assessment in math over a 7 month period; 2) change in intensity of concern about standards-based instruction and assessment over a 7

month period; and 3) changes in frequency of use of standards-based instruction and assessment strategies in mathematics classes.

As described in Chapter 2, a concern is “the composite representation of feelings, preoccupation, thought, and consideration given to a particular issue or task [such as standards-based instruction and assessment in math” (Hall & Hord, 1987, p. 9).

According to the Concerns-Based Adoption Model (CBAM) of Change, individuals facing an innovation experience a series of concerns as they became more familiar with and adept in the use of an innovation. The concerns are divided into 7 stages, or levels, and it is necessary for early stage concerns to either be resolved or at least reduced in intensity before later, more mature concerns, can emerge or increase in intensity. In contrast to the level of concern, intensity of concern refers to the overall strength of the feelings, preoccupation, thought, or consideration one gives to a particular issue or task. Intensity of concern cannot be interpreted directly; a high level of concern is not necessarily good, and a low intensity of concern is not necessarily bad. For example, a teacher who has incorporated standards-based instruction and assessment into his/her repertoire and is completely comfortable with the strategies associated with it might report low levels of concern with standards-based instruction and assessment. By the same token, a teacher who is barely aware that standards-based instruction and assessment exists will also demonstrate low levels of concern about this innovation.

In addition, individuals at very different levels of concern can view an educational innovation with the same intensity. For example, the overall intensity of concern is generally in the same range for a person at a high level 6 (refocusing) as it is for a person

at a high level 1 (informational). Both individuals are experiencing high levels of preoccupation or mental arousal about standards-based instruction and assessment in math. However, the level 1 person is intensely concerned about the general characteristics, effects, and requirements for using standards-based instruction and assessment, while the level 6 is intensely concerned about the universal benefits and possible replacements for this innovation.

In this study, the variable, *changes in levels of concern about standards-based instruction and assessment*, was used to describe the shift over time from one level of concern to another in the Concerns-Based Adoption Model (CBAM) of Change. The levels of concern in the CBAM model are (in order from low to high): 0: awareness; 1: informational; 2: personal; 3: management; 4: consequence; 5: collaboration; and 6: refocusing.

The variable, *change in intensity of concern*, was developed to describe changes in overall intensity of concern (negative or positive) during the 7 month staff development project. This variable differed from the change in level of concern variable in that it represented shifts in overall intensity of concern over time, rather than movement from one particular level of concern to another.

The variable, *changes in frequency of use* of standards-based instruction and assessment strategies, designated the degree to which teachers had increased or decreased their use of standards-based techniques over the course of the project. Techniques included: aligning instruction and assessment with standards, requiring students to explain how they arrived at their answers, employing rubrics, encouraging students to

work in groups, using manipulatives in math class, keeping written records of student performance, observing more than one sample of student performance before grading, among others.

Population

Teachers

The study population consisted of 205 Rhode Island elementary school teachers engaged in the second year of a staff development project aimed at increasing their familiarity with and skill at implementing standards-based instruction and assessment in mathematics.

This study focused on the second year (academic year 1999/2000) of a state-funded staff development initiative designed to change curriculum, instruction, and assessment in substantive ways. Goals of this project included the following:

- Teachers will report increased understanding of standards-driven curriculum, instruction, and assessment
- Teachers will use, develop, and implement standards-driven units/curriculum
- Teachers will identify gaps in their own knowledge and training and identify resources to fill those gaps

During the 1999/2000 academic year, the staff development project consisted of four essential components: 1) a series of three 1 ½ day lead teacher workshops focusing on action research and standards-based education; 2) local consultants who worked closely with each school; 3) implementation of action research projects at the school level; and 4) a commitment by participants to spend 25 hours on staff development at the

school level.

Fifty-eight K-6 teachers in the population (i.e., 28%) were “lead” teachers, meaning that they had volunteered to attend the three workshops. Lead teachers were also responsible for bringing work and information back to their schools, supporting other teachers involved in the project, and conducting an action research project. In addition, lead teachers were expected to devote 25 hours to in-school activities that they designed and/or requested, in order to fill their own and other teachers’ knowledge and training gaps related to standards-based instruction and assessment. Each lead teacher received an honorarium of \$1000 for the 1999/2000 academic year.

One hundred forty-seven K-6 teachers in the population (i.e., 72%) were “participating” teachers. This designation indicated that their involvement in the staff development program consisted of 25 hours of staff development activities and completion of an action research project. Each participating teacher received an honorarium of \$500 for the 1999/2000 academic year and was not required to attend any of the 1 ½ day lead teacher workshops focusing on action research and standards-based education.

Students

The student population consisted of approximately 850 students in grades 2 through 5 in the classes of teachers engaged in the staff development project.

Subjects

Teachers

The study focused on a sample of 89 Rhode Island elementary teachers in 11 schools, all of whom were participating in the staff development project. This sample corresponds to 43% of the population from which the sample was drawn.

Eighty-five of the 89 teachers in the sample taught at regular elementary schools, while 4 teachers taught elementary grades at a school for the deaf serving students from pre-school through grade 12. Furthermore, 16 teachers taught in 3 rural schools, while the remaining teachers were employed in 8 urban schools. The sizes of the schools in which subjects taught ranged from 122 students to 817 students; however, the majority of study subjects (81/89) worked in schools with 300 or more students.

Thirty-three subjects were lead teachers in the staff development project, while 56 were participating teachers. Consequently, there were proportionately more lead teachers in the sample (37%) than in the population (28%). While lead teachers in this sample attended more workshops ($p < .05$) and meetings ($p < .05$) than participating teachers, they did not differ significantly from participating teachers on any other staff development, contextual, or personal independent variables. However, the large proportion of lead teachers in the sample may indicate that the results of this study are not generalizable to the population of teachers.

Students

The study also focused on a sample of 607 second through fifth grade students who provided data regarding their perceptions of the degree to which standards-based mathematics was implemented in their classrooms.

Methods

Teacher survey: pre

In the fall of 1999, teachers completed a 108-item survey designed to collect data on the following: a) level of concern about standards-based instruction and assessment in mathematics; b) intensity of concern about standards-based instruction and assessment in mathematics; c) teacher role in the staff development project (i.e., lead vs. participating teacher), d) total number of years teaching experience; e) teachers' sense of self-efficacy; f) quality of school-wide communication; g) degree of principal directiveness; h) degree of principal supportiveness; i) quality of school/community relations; and j) quality of school's organizational climate. The number of survey items used to measure each variable is shown in Table 1 below. A copy of the teacher pre-survey is contained in Appendix 1.

To measure changes in teachers' levels and intensity of concern toward standards-based instruction and assessment in math, the 35 items comprising the Stages of Concern (SoC) about the Innovation Questionnaire were included among the 108 items in the pre-survey. The 35-item SoC Questionnaire, which was developed by The Research and Development Center for Teacher Education at the University of Texas Austin in 1976,

comprised seven scales corresponding to the seven levels, or stages, of concern about an innovation identified in the Concerns-Based Adoption Model (CBAM) of change.

Each of the 35 SoC items consisted of a statement of concern about standards-based instruction and assessment in math. Respondents were instructed to rate each item

Table 1

Number of survey items per variable

Variable	Number of items
Level of Concern/Intensity of Concern	35
Teacher Role	1
Size of School	1
Number of Years Teaching	1
General Self-Efficacy	16
Quality of School-Wide Communication	9
Degree of Principal Directiveness	7
Degree of Principal Supportiveness	9
Quality of School/Community Relations	7
Quality of Organizational Climate	9
Frequency of Use of Standards-Based Instruction and Assessment in Math	13
Total Items	108

on a 0 to 5 Likert scale according to how true it is that the item described a concern felt by the individual at the that time (1=very untrue of me at this time, 5=very true of me at

this time). The “0” at the end of the scale was recommended for marking items that were completely irrelevant.

Table 2 displays the alpha coefficients of internal consistency for each of the seven SoC scales, as presented in the Manual for the Use of the SoC Questionnaire (Hall, George, & Rutherford, 1977, p. 10). These alpha coefficients, which range from .64 to .83, are quite adequate, given the fact that alpha coefficients higher than .60 are generally considered sufficient for questionnaire studies. Coefficients of internal consistency on the seven SoC scales for this study sample were calculated and are reported in Chapter 4.

Table 2

Coefficients of Internal Reliability for the SoC Questionnaire, N=830

Stage	0	1	2	3	4	5	6
Alphas	.64	.78	.83	.75	.76	.82	.71

Teachers’ sense of self-efficacy was measured via 16 items from the 17-item General Self-Efficacy Scale developed by Sherer, Maddux, Mercandante, Prentice-Dunn, Jacobs, and Rogers (1982). The only item from the original General Self-Efficacy Scale that was not used in this study was: “I do not seem capable of dealing with most of the problems that come up in life.” It was feared that this item would be found offensive by teachers and was thus omitted.

Each of the 16 general self-efficacy items used consisted of a statement and instructions to teachers to rate each statement on a 1 to 4 Likert scale according to how typical it was of them (1=very atypical, 4=very typical). The alpha reliability for the

original version of the General Self-Efficacy Scale was reported by Sherer et al. (1982) as .86. The coefficient of internal consistency on the General Self-Efficacy Scale for this study sample was calculated and is reported in Chapter 4.

Survey items corresponding to the quality of school-wide communication (9 items), quality of school/community relations (7 items), and quality of school organizational climate (9 items) were selected from appropriate subscales of the Troubleshooting Checklist for School-Based Settings (Manning, 1976), a survey instrument that focuses on the strengths and weaknesses of a school's environment in relation to the adoption and implementation of educational innovations. Alpha reliabilities for these three scales were reported by Manning (1976) as .89 (quality of school-wide communication), .82 (quality of school/community relations), and .87 (quality of school organizational climate). Coefficients of internal consistency on each of these scales for this study sample were calculated and are reported in Chapter 4.

The degree of principal directiveness was measured through teachers' responses to 7 items from the Principal Directiveness Subscale of the Organizational Climate Description Questionnaire (Halpin, 1963). The degree of principal supportiveness was assessed through teachers' responses to 9 items from the Principal Supportiveness Subscale of the Organizational Climate Description Questionnaire (Halpin, 1963). Each of these items consisted of a statement and instructions to teachers to rate each statement on a 1 to 4 Likert scale according to how typical it was of them or their school (1=very atypical, 4=very typical). Stallings and Mohlman (1981) reported alpha reliabilities for

these Principal Directiveness and Principal Supportiveness scales as .76 and .88, respectively.

A 14-item scale designed to measure teachers' use of standards-based instruction and assessment practices in their math classes was developed by the researcher. Items were drawn from literature related to standards-based instruction and assessment. Each item described an instructional or assessment strategy. Each item was accompanied by instructions to rate each strategy on a 1 to 5 scale according to how often they employed each strategy in their math classroom. Ratings for this scale were: 1: less than one time per quarter; 2: one time per quarter; 3: 2-3 times per month; 4: 1-4 times per week; and 5: 5 or more times per week. Coefficients of internal consistency for this sample were .90 for the pre-survey and .79 for the post-survey.

Finally, two survey items were developed to identify teachers' roles in the project (i.e., lead teacher versus participating teacher) and the number of years in which they had been teaching: 1 year or less, 2-5 years, 6-10 years, 11-19 years, and 20 or more years.

Teacher survey: post

A post-survey consisting of the 35 SoC items and the 14 items assessing teachers' utilization of standards-based instruction and assessment strategies in mathematics was developed to measure changes in teachers' level and intensity of concern about and use of standards-based instruction and assessment since earlier in the school year.

Professional development activity logs

Data concerning the types of professional development activities in which teachers engaged (during their 25 "free" hours) and the amount of time they spent on each type of

activity were gathered by means of a Professional Development Activity Log that was completed by each teacher. Each log consisted of blocks corresponding to weeks in the school year. Each block was divided into two columns. In one column, respondents were instructed to describe the activity in which they had participated. In the second column, respondents were to indicate the amount of time spent on each activity. A copy of the Professional Development Activity Log is contained in Appendix 2.

Student surveys

A student survey was designed by the researcher to explore the degree to which students perceive that their teachers are using typical standards-based instruction and assessment activities/tasks in their math classes. The student questionnaire consisted of 16 statements describing the types of math activities in which students engaged and the math tasks that teachers assigned. Students were instructed to rate each item on a 1 to 4 Likert scale according to how often they engaged in these activities or completed the indicated tasks (1=never, 2=sometimes, 3=most of the time, 4=always). The reliability on the student questionnaire for this sample was .59. The student survey is contained in Appendix 3.

Teacher Interviews

An important component of this study was qualitative in nature. Teachers' perceptions of the impact of various staff development activities on their classroom practices (in math) were explored via 132 formal interviews with 71 teachers engaged in the project between October 1999 and June 2000. Key interview questions and probes were developed by the researcher to generate data that would identify the various types of

staff development activities that teachers felt were affecting their beliefs and practices to the greatest degree. Interview questions and probes can be found in Appendix 4.

Collection of Data

Teacher survey: pre

In December 1999, pre-surveys were mailed directly to 205 teachers at their schools. Each teacher was assigned a code, and each survey included a computer scanning sheet with the appropriate code already filled in. A cover letter attached to the survey informed teachers of the purpose of the survey and the value of their participation in the project. To encourage prompt completion of the survey, stamped envelopes (addressed to the researcher) were included in the mailing. In addition, teachers were informed that those who returned the survey by the deadline would be entered in a drawing for 3 gift certificates to local department stores.

To follow up with teachers who did not respond to the questionnaire, a second mailing was conducted in March 2000. This time, survey packets were sent directly to school principals, who were informed of the purpose of the study and asked to distribute surveys to the teachers.

Post-surveys were mailed to teachers (via their principals) in May 2000. To facilitate the tracking of changes in teachers' concerns, each survey was coded with each teacher's unique code. As with the pre-survey, each teacher was sent a questionnaire, a computer scanning answer sheet, a cover letter explaining the purpose of the survey, and a stamped envelope (addressed to the researcher) with which to mail the completed survey.

Professional development activity logs

Professional development activity logs were distributed to all lead teachers in October 1999. One log covered the time period from October 1999 through March 2000, while a second log was to be used for activities conducted from April 2000 through May 2000. The logs were explained at the time of distribution, and lead teachers were asked to distribute logs to participating teachers at their schools. Prior to the March and May 2000 lead teacher workshops, lead teachers collected logs from participating teachers. They subsequently brought their own and participating teachers' logs to the lead teacher workshops, where they were collected by project staff.

Teacher interviews

One hundred thirty-two interviews with 71 teachers were conducted between October 1999 and June 2000 on site at 5 of the 11 schools involved in this study. While the majority of the interviews were conducted one-on-one, a few interviews were held with groups of 2-3 teachers. Since these interviews fit a qualitative approach to research, the exact wording and sequence of questions varied. However, the majority of prompts focused on the staff development activities and contextual factors that were facilitating and/or hindering teachers' adoption and implementation of standards-based teaching practices. Perceived changes in teaching practice and student performance were also noted.

Interviewees were encouraged to tell their own stories in their own words; these words were subsequently used by the researcher to format additional questions and

probes. The researcher transcribed the content of these interviews on a laptop computer as interviewees spoke.

It was the assumption of the researcher that the views expressed by teachers in the five schools in which interviews took place would generalize to teachers in the remaining six schools in which interviews were not held. In fact, the five “interview schools” were not selected because they were qualitatively different from the others. Rather, budgetary constraints precluded interviewing teachers from every school in the initiative, and a conscious attempt was therefore made to interview teachers from at least one school per district. Furthermore, statistical analyses revealed that teachers in the interview schools did not differ significantly from teachers in non-interview schools on any of the independent or dependent variables ($p < .05$).

Student surveys

The student questionnaire was administered in May 2000 to a random sample of 40 classes of grade 2-6 taught by teachers in the study sample. The decision was made to distribute the student questionnaire to a random sample of 40 teachers’ classes in order to maximize the probability of receiving completed surveys from 28 classes, the sample size necessary to detect a significant correlation at $\alpha = .05$ when the population r is large, i.e., at least .50 (Cohen, 1992).

Using a table of random numbers and teacher codes, the researcher randomly selected a sample of 40 grade 2-5 teachers to whom student surveys would be sent.

Student surveys were mailed to these 40 teachers in May 2000. Teachers were asked to

administer the 10 minute survey in their math class and return it to the researcher within 2 weeks in the stamped envelope that was provided.

Treatment of Data

Survey data were entered into the Statistical Package for the Social Sciences (SPSS) Version 7.5.

Teacher survey data: pre

Based on their responses to the 35 SoC items embedded in the 108-item pre-survey, teachers received a total score for each of the 7 scales contained in the SoC questionnaire. These seven scale totals were summed to obtain an intensity of concern (toward standards-based instruction and assessment) pre-score for each teacher. Each total intensity of concern score represented “a straightforward indication of the intensity of concern” (Hall & Loucks, 1977, p. 49). Intensity scores ranged from 35 to 175, with low scores indicating low intensity of concern and high scores indicating high intensity of concern.

In addition, teachers were assigned to the stage of concern that most aptly described them; this was represented as a level of concern pre-score for each teacher. Stage of concern scores ranged from 0 to 6. To obtain this level of concern pre-score, each teacher received a “raw score” (i.e., the sum of the responses to the five statements on that scale) for each of the 7 Stages of Concern scales. The level of the highest stage score was subsequently utilized to designate each teacher’s pre-level of concern. The higher the stage score, the further along teachers were in adapting to and innovating standards-based instruction and assessment in mathematics.

Based on their responses to the use of standards-based instruction/assessment scale, teachers received a frequency of use (of standards-based instruction and assessment techniques) score. Frequency of use scores ranged from 14 (infrequent use of standards-based instruction and assessment) to 70 (frequent use of standards-based instruction and assessment).

Derived from their responses to the self-efficacy, degree of principal directiveness, degree of principal supportiveness, quality of school-wide communication, quality of school/community relations, and quality of school organizational climate sub-scales, teachers also received scores for each of these additional sub-scales. Each person's score on these variables was calculated as the average of all of his/her responses to the items on that particular scale. Average scores on each of these variables ranged from 1 to 5, with 1 indicating a low degree of each variable.

Missing data on these scales was handled in the following manner. If subjects responded to fewer than half of the items on each scale, they received a missing score on that variable. If, however, subjects responded to half or more of the items on a particular scale, they received a scale score on that variable that was the mean of the responses to the non-missing items.

Teacher survey: post

Based on their responses to the 35 SoC items embedded in the 49-item post-survey, each teacher received an intensity of concern post-score and a level of concern post-score (toward standards-based instruction and assessment). Teachers also received a frequency of use (of standards-based instruction and assessment techniques) post-score,

based on their responses to the remaining 14 items in the survey. These scores were computed exactly as they were for the pre-survey.

Changes in level of concern were obtained by regressing teachers' post-SoC scores on pre-SoC scores. From this procedure, a deviation or residual gain score was obtained for each teacher in the sample; this residual gain score was termed *change in level of concern*. The change in level of concern variable was subsequently used as a measure of gain in concern above and beyond what was predicted by the pre-SoC level of concern alone.

The *change in intensity of concern* variable was calculated in a similar fashion. Based on their responses to the pre-SoC sub-scale, each teacher was assigned a total concern score (0-245). The same procedure was conducted following teachers' completion of the post-SoC questionnaire. The change in intensity of concern variable was a residual gain score obtained by regressing teachers' post-total concern scores on pre-total concern scores. As a result, the change in intensity of concern variable was used as a measure of gain in intensity of concern above and beyond what was predicted by the pre-total concern score alone. Similarly, the change in frequency of use variable was obtained to represent residual gains in frequency of use of standards-based instruction and assessment techniques.

Professional development activity logs

Logs were collected and content-analyzed in April 2000 and June 2000. Based on a careful review of teachers' logs, codes were developed to correspond with the major staff development strategies used by study participants. In addition, the amount of time

each study subject engaged in each staff development strategy was noted. These codes (*action research, workshops, meetings, work with educational consultants, carrying out standards-based activities in the classroom, independent research, professional reading, meetings/discussions with colleagues, curriculum development, and examining student work*) were entered as variables into SPSS. Finally, the total amount of time each teacher spent participating in each form of staff development was entered as the “score” on each variable.

Interview data

Interview data were coded according to categories that emerged from the data as well as pre-set categories (developed by the researcher) that had been identified through teachers’ professional development activity logs. The qualitative software program, HyperRESEARCH, was utilized to simplify this task. Reports were run to organize teacher data.

Student survey data

Based on their responses to the student questionnaire, each student received a total score on the student scale. A mean score for each class was then computed. This mean class score was used as the value for the student score variable that was subsequently created for each teacher.

Statistical methods

For each of the three hypotheses being tested, stepwise regression analyses were performed to explore the utility of using school size, teacher role, teacher self-efficacy, total number of years teaching experience, quality of school-wide communication, degree

of principal directiveness, degree of principal supportiveness, quality of school/community relations, quality of organizational climate, type of activity engaged in, and intensity of activity to predict *residual gains* in the dependent variable (level of concern, intensity of concern, or frequency of use of standards-based instruction and assessment).

According to statisticians (Thompson, 1995; Licht, 1997; Pedhazur, 1997; Tabachnik & Fidell, 1996), computer statistical packages such as SPSS incorrectly compute degrees of freedom in stepwise regression analyses, resulting in Type I error rates in excess of the alpha level selected by the researcher. This is due to the fact that, at each step, predictors are selected by looking at the results for all of the predictor variables not entered. However, potential independent variables do not enter the final regression equation, and the reported F is biased because it does reflect any degrees of freedom for the variables that were “consulted” but not selected.

When inference in stepwise multiple regression is desired, most researchers agree that adjustments are necessary for reducing the occurrence of Type I errors. While there is no agreed upon method for doing so, there is consensus that conclusions made from procedures involving stepwise regression with adjustments be limited to prediction, as opposed to explanation.

“There is no single recommended or agreed upon approach for tests of significance in predictor-selection procedures. Whatever approach is followed, it is important that its use be limited to the case of prediction. Adverse effects of biased Type I errors pale in comparison with the deleterious consequences of using predictor-selection procedures for explanatory purposes” (Pedhazur, 1997, p. 219).

The statistical adjustment utilized in this study is credited to Wilkinson and Dallal (1981), who developed tables for critical R^2 when forward selection procedures are used for statistical addition of variables and the selection stops when the F-to-enter for the next variable falls below some preset value. Wilkinson and Dallal's tables identify how large R^2 must be to be statistically significant at a given alpha level, given N, k, and F, where N is sample size, k is the number of potential predictor variables, and F is the minimum F-to-enter specified by the researcher. Using information from Wilkinson and Dallal's tables, adjustments in tests of statistical significance results were made, thus reducing the risk of Type I errors.

Using traditional product term analysis, two-way interactions among the independent variables in the final regression equations were explored. This was accomplished via a second set of direct entry regression analyses in which the variables from the final regression equation and the interaction terms were regressed on the outcome variable (either level of concern, intensity of concern, or frequency of use of standards-based instruction and assessment). These results were examined for the presence of an interaction effect, the strength of any significant interaction effects, and the nature of any significant interactions. Finally, main effects were interpreted.

To check the concurrent validity of teachers' reported use of standards-based instruction and assessment, correlations were run between teachers' raw scores on the post-version of the use of standards-based instruction and assessment scale and the mean student-reported score of the same construct. Interview data were used to complement statistical findings and lend concrete examples and explanations to quantitative results.

Limitations

The following limitations of this study have been recognized:

1. The amount of time between pre- and post-administrations of teacher surveys was very short. This time period may not have been sufficient for changes in concern and use to occur. In addition, this study focuses on staff development and contextual features during the second year of a change initiative. As such, it is possible that changes in teachers' concerns and practices took place in year one of the project, before this study began. The nature of this study also does not provide a full picture of the degree of long term change that may have occurred in teachers' concerns and classroom practices.

2. Teacher concerns, teacher and school characteristics, and staff development activities were measured by self-report through use of a survey and a professional development log. This data collection approach carried with it the risk that respondents were not entirely truthful in their responses. This could stem from social desirability (e.g., teachers may have wanted to appear very self-efficacious or they may have indicated that they participated in more staff development than they actually did) or even fear of repercussion (e.g., some teachers may not have described their principal truthfully for fear that s/he would see their responses). As with all self-report, one must keep in mind that a person's account of his/her activities, attitudes, or perceptions is not entirely neutral and does not necessarily provide the researcher with the whole or final picture (Keeves, 1997).

3. The student survey was not administered to kindergarten and first grade students. As a result, findings regarding the degree of correlation between teachers'

reported use of standards-based instruction and assessment and their students' perception of teachers' use of standards-based instruction and assessment may not be generalizable to these grade levels.

4. Because the study sample (89) represented only 43% of the population (205), it is possible that the teachers in the sample are not representative of the teachers in the population. Considering the demands put on teachers in this study (completing 2 surveys and keeping a weekly log), one might even surmise that the teachers in the sample were particularly motivated or hard-working, compared to the teachers who did not complete the surveys or the log. Future research will need to address the issue of whether the results of this study apply to all teachers or just the motivated group.

5. A sample size of 89 is very small for stepwise regression analysis. The use of a small sample in stepwise regression carries with it the risk that the solutions that are yielded will be spurious or will not generalize beyond that particular sample. Analyses with a small sample may also lack the power necessary to detect differences that in fact exist.

However, the reader should keep in mind that this is an exploratory study. Staff development activities and contextual features have never before been examined as a means of predicting change in teachers' attitudes or classroom practices. As such, the results of the stepwise regression analyses performed will need to be explored further in future research.

Summary

The general characteristics of the teacher and student sample have been described in this chapter. Descriptions of survey instruments, the independent and dependent variables, the methods used in collecting and treating the data and testing the hypotheses were also provided. The next chapter discusses the results of the analyses described in this chapter.

CHAPTER 4

Results

The results presented in this chapter are organized into six main sections: (1) response rates and sample size; (2) descriptive statistics related to the independent and dependent variables; (3) analysis of the relationships among independent variables and residual gains in teachers' level of concern toward standards-based instruction and assessment in mathematics; (4) analysis of the relationships among the independent variables and residual gains in teachers' intensity of concern toward standards-based instruction and assessment in mathematics; (5) analysis of the relationships among the independent variables and residual gains in teachers' reported use of standards-based instruction and assessment in mathematics; (6) analysis of teacher interviews; and (7) summary of results.

Response Rates and Sample Size

In the December 1999, pre-surveys were mailed directly to 205 teachers at their schools. By February 2000, teacher response rate was 45%. In an attempt to achieve a better response, a second mailing was conducted in March 2000. Forty-five additional surveys were returned, bringing the overall response rate to 70% (n=143).

The response rate for the post-survey (administered in June 2000) was 56% (n=120). However, only 102 of the teachers responding to the post-survey had also completed their pre-survey, lowering the effective sample size at that point to 102.

In all, 182 teachers completed and submitted their professional development activity logs. Unfortunately, only 89 of these 182 teachers also completed both the pre-

and the post-survey. As a result, the final sample size for the statistical component of this study was reduced to 89.

Descriptive Statistics

Analyses of the surveys and professional development activity log allow for a presentation of descriptive statistics for the independent and dependent variables in this study. This description is presented in this section.

Independent Variables

Teacher Role. Thirty-three subjects were lead teachers in the staff development project, while 56 were participating teachers.

Size of Schools in which Teachers Work. The sizes of the schools in which teachers taught ranged from 122 students to 817 students; however, the majority of study subjects (81/89) worked in schools with 300 or more students. Table 3 presents a breakdown of school size by number of study subjects.

Number of Years Teaching Experience. Seventy-nine percent of the teachers in the study sample indicated how many years they have been teaching. On the whole, these individuals had been teaching for quite a few years. Slightly more than 22% of teachers reported that they had been teaching twenty or more years, while just over 19% of teachers had been teaching between 11 and 19 years. Almost 17% of the sample had taught for 6 to 10 years, while almost 20 % of teachers in the sample had taught five years or fewer. Nineteen teachers in the sample (21.3%) declined to report how many years they had been teaching, suggesting that teachers found questions regarding the length of their teaching experience to be threatening or offensive.

Table 3

School Sizes

School	Student population	Teachers in study sample (Percent of sample)	
1.	138 students	4	(4.5%)
2.	302 students	4	(4.5%)
3.	326 students	12	(13.5%)
4.	426 students	8	(9.0%)
5.	437 students	8	(9.0%)
6.	477 students	8	(9.0%)
7.	570 students	5	(5.6%)
8.	594 students	15	(16.9%)
9.	733 students	9	(10.1%)
10.	817 students	12	(13.5%)
TOTAL		89	(100%)

Teachers' Sense of Self-Efficacy. Table 4 displays descriptive statistics related to teachers' reported self-efficacy (i.e., their belief in their ability to perform a variety of tasks, as well as their willingness to initiate and persist in the completion of a task).

As shown in Table 4, all study subjects reported a rather high degree of self-efficacy; teachers' mean self-efficacy score was 3.51 on a scale of 1 (low self-efficacy) to 4 (high self-efficacy). Teachers in the study sample clearly felt that they were capable of performing a number of tasks and seeing them through to completion. The coefficient of internal consistency for the General Self-Efficacy scale for this sample was .76.

Table 4

Descriptive Statistics: General Self-Efficacy, School-Wide Communication, Principal Directiveness, Principal Supportiveness, School/Community Relations, and Organizational Climate (N=89)

	N	Minimum	Maximum	Mean	SD	Cronbach's α
General Self-Efficacy	89	2.69	4.00	3.51	.32	.76
Quality of School-Wide Communication	89	1.33	4.00	2.83	.61	.84
Degree of Principal Directiveness	89	1.00	4.00	2.52	.56	.58
Degree of Principal Supportiveness	89	1.67	4.00	2.97	.61	.88
Quality of School/Community Relations	89	1.86	4.00	2.92	.43	.47
Quality of Organiz. Climate	89	1.89	3.67	2.75	.42	.62

Quality of School-Wide Communication. As shown in Table 4, the quality of school-wide communication reported by teachers was somewhat variable, with average scores ranging from 1.33 to 4.00 on a scale of 1 (low quality communication) to 4 (high quality communication). Overall, the quality of school-wide communication was average, with a mean of 2.83, meaning that teachers perceived that a fair amount of face-to-face communication among school and project staff took place at their school. However, these results suggested that there was room for improvement; high quality

school-wide communication was not yet the norm in most schools. The coefficient of internal consistency for the Quality of School-Wide Communication scale for this sample was .84.

Degree of Principal Directiveness. As shown in Table 4, the overall degree of principal directiveness was lower than the other school level scores, with a mean of 2.52 on a scale of 1 (low degree of principal directiveness) to 4 (high degree of principal directiveness). An average score of 2.52 on the principal directiveness scale suggests that, across schools, teachers found it neither typical nor atypical for principals to make all scheduling decisions, control agendas and activities, and allow their rules to be questioned. This suggested that principals were not very consistent with respect to the degree of directiveness they exhibited. The coefficient of internal consistency for the Degree of Principal Directiveness scale for this sample was .58.

Degree of Principal Supportiveness. Data in Table 4 suggest that principals at study participants' schools were more supportive than they were directive ($p < .001$), with a mean directiveness rating of 2.97 on a scale of 1 (low degree of principal supportiveness) to 4 (high degree of principal supportiveness). In addition, no teachers in the study rated their principals a "1" on this scale, indicating that all principals were seen as at least somewhat supportive. These findings suggested that principals in teachers' schools were typically willing to work with teachers and share their educational knowledge with them. These principals were also perceived by teachers as communicating clearly and offering helpful, constructive criticism. The coefficient of

internal consistency for the Degree of Principal Supportiveness scale for this sample was .88.

Quality of School/Community Relations. Overall, the quality of school/community relations was fairly high, with a mean of 2.92 on a scale of 1 (low quality school/community relations) to 4 (high quality school/community relations). This mean rating indicated that it was somewhat typical for parents and community members to be aware of and involved in innovations and school events and to offer their opinions to school administrators. Descriptive statistics for this variable are displayed in Table 4. The coefficient of internal consistency for the Quality of School/Community Relations scale for this sample was .47.

Quality of Organizational Climate. As shown in Table 4, the quality of school organizational climate reported by teachers was similar to the other school-level scores, with mean scores ranging from 1.89 to 3.67 on a scale of 1 (low quality climate) to 4 (high quality climate). Interestingly, however, not a single study participant rated his/her school climate a 4 (the highest rating), meaning that none of the teachers in the sample thought the organizational climate of his/her school was ideal. Overall, the quality of school-wide communication was average, with a mean of 2.75. In other words, teachers felt that their schools were somewhat open to innovation and that the channels and procedures for dissemination of information were average. Given the average nature of these ratings (i.e., neither very high nor very low), it was also likely that a few isolated subsystems of teachers existed within the study schools and that not all teachers were involved or interested in the change initiative that was being implemented. The

coefficient of internal consistency for the Quality of Organizational Climate scale for this sample was .62.

Types of Staff Development Activities. The principle staff development activities reported by teachers in their professional development activity logs were: action research, workshops, meetings, work with educational consultants, carrying out standards-based activities in the classroom, independent research, professional reading, meetings/discussions with colleagues, curriculum development, and examining student work. Table 5 presents descriptive statistics on the number of hours teachers reported that they devoted to each activity.

It is interesting to note that the professional development activity engaged in the most by teachers was not a traditional training activity, such as attendance at a workshop. Rather, teachers reported that the act of implementing new standards-based curriculum or experimenting with different types of standards-based activities was in itself a professional development activity. This represents a radical shift in the notion of staff development, as teachers appeared to recognize that staff development (1) is closely connected with classroom practice; (2) can take place within the four walls of one's classroom; and (3) often occurs during the course of the school day.

After standards-based classroom activities (mean=17.22 hours), teachers reported that they participated in meetings (mean=12.82 hours), professional reading (mean=10.67 hours), and workshops (mean=8.24 hours) most often from October 1999 to May 2000. Finally, they indicated that they did very little action research, curriculum development,

examination of student work (beyond the normal tasks of grading papers and tests), or independent research, with means of 2 hours or less during that seven-month period.

Table 5

Staff Development Activities Engaged in by Study Participants

	N	Minimum # Hours	Maximum # Hours	Mean # Hours	SD
Standards-Based Classroom Activities	89	0	211.5	17.2	31.0
Action Research	89	0	18.5	2.1	3.9
Work with Consultant	89	0	14.0	3.2	3.7
Curriculum Development	89	0	18.0	1.5	3.1
Meetings	89	0	86.3	12.8	15.3
Professional Reading	89	0	71.8	9.5	10.7
Independent Research	89	0	21.5	1.0	3.2
Examining Student Work	89	0	21.0	1.2	3.4
Workshops	89	0	51.5	8.2	12.0

Dependent Variables

Change in Level of Concern about Standards-Based Instruction and Assessment.

Analyses revealed that study teachers were very homogenous with regards to their pre- and post-levels of concern. As displayed in Table 6, more than one third of teachers (n=33) were at a level 1 stage of concern (awareness) both at the time of the pre-survey and the post-survey. In addition, nearly all of the teachers in the sample scored between levels 1 and 4 at pre-test (n=71) and post-test (n=78). In addition, little variability was exhibited within each scale, indicating that teachers tended to respond similarly to the questions on the SoC surveys. As evidenced by the standard deviations shown in Table 6, teachers' scores cluster very tightly around the mean of each scale.

It is obvious from Table 6 that the coefficients of internal consistency obtained for each 5-item SoC scale at pre-test and post-test are extremely low. In fact, they range from just .22 to .45. This is no doubt due to the homogeneity of the teachers in this sample. When a group is homogeneous, a restriction of range results, true score variance is reduced, and low reliability coefficients are yielded.

Table 7 displays mean pre-levels of concern, post-levels of concern, and residual gain scores in level of concern for teachers who participated in this study. As evidenced in Table 7, there was almost no change in overall level of concern during the seven-month study period. In fact, teachers' average level of concern dropped by just two one-hundredths of a level.¹ Moreover, paired sample t-tests revealed that there was

¹ NB: The very small mean residual gain in concern score of virtually zero (3.37×10^{-16}) cannot be interpreted directly. The mean value of any set of residuals is always zero when it is calculated on the same n pairs that produced two sets of parallel measurements (Glass & Hopkins, 1996).

no significant difference between teachers' pre- and post-levels of concern about standards-based instruction and assessment ($p=.36$). This information

Table 6

Level of Concern Results from Pre- and Post-SoC Surveys

Pre-SoC Survey					Post-SoC Survey				
SoC Scale	Mean (Range 5-35)	SD	Alpha Reliability	# Teachers at This Level	SoC Scale	Mean (Range 5-35)	SD	Alpha Reliability	# Teachers at This Level
Level 0	15.37	2.75	.34	2	Level 0	14.88	2.67	.28	1
Level 1	19.07	2.79	.27	33	Level 1	18.52	2.95	.34	33
Level 2	17.72	2.98	.35	17	Level 2	16.85	3.21	.35	12
Level 3	16.48	2.82	.29	10	Level 3	16.51	2.76	.22	19
Level 4	16.73	3.21	.44	11	Level 4	16.00	3.17	.34	14
Level 5	16.78	2.82	.45	9	Level 5	16.20	2.97	.45	5
Level 6	16.90	2.64	.29	7	Level 6	16.52	2.88	.39	5

suggests one of two things: either 7 months is not long enough to detect growth in broad levels of concern about standards-based instruction and assessment or staff development and/or school and personal conditions were insufficient to foster statistically significant changes in levels of concern.

Table 7
Pre, Post, and Residual Gains in Level of Concern

	N	Minimum	Maximum	Mean	SD
Pre-Level of Concern	89	0	6	2.48	1.70
Post-Level of Concern	89	0	6	2.46	1.54
Residual Gain in Level of Concern	89	-2.78	3.68	3.37×10^{-16}	1.53
Valid N (listwise)	89				

Change in Intensity of Concern about Standards-Based Instruction and Assessment. Table 8 displays mean pre-intensity of concern, post-intensity of concern, and residual gain scores in intensity of concern for teachers who participated in this study.

On a scale of 35 (low intensity of concern) to 175 (high intensity of concern), teachers' mean intensity scores shifted from 126.69 to 121.49 over the course of the research project. This apparent decrease in intensity could indicate one of two things:

either teachers became more comfortable/familiar with standards-based instruction and assessment in math (and hence their concerns became less intense) or they lost some

Table 8

Pre, Post, and Residual Gains in Intensity of Concern

	N	Minimum	Maximum	Mean	SD
Pre-Intensity of Concern	89	73.50	180.00	126.69	20.03
Post-Intensity of Concern	89	57.50	159.50	121.49	20.82
Residual Gain in Intensity of Concern	89	-66.04	36.77	1.50×10^{-15}	17.81
Valid N (listwise)	89				

interest in standards-based instruction and assessment. (As in the previous discussion, the mean residual gain in intensity of concern score of approximately zero (1.50×10^{-15}) should not be interpreted directly.) It is not possible to make a quality judgment regarding the nature of the intensity of concern based on the intensity of concern score alone. Rather, this score simply reflects overall intensity of concern.

Paired sample t-tests revealed that teachers' initial intensity of concern was statistically significantly different from their post-intensity of concern about standards-based instruction and assessment ($p=.02$). However, it should be kept in mind that the actual gain in level of use (beyond what is predictable from the pre-survey alone) is less than one point (1.50×10^{-15}) on a 140-point scale, and not -5.2 points, which

would be obtained if one simply subtracted the mean pre-intensity score (126.69) from the mean post-intensity score (121.49). This is the advantage of using residual gain scores, which measure how much change has occurred unrelated to initial performance. The coefficients of internal consistency for intensity of concern based on this sample were .82 for the pre-survey and .81 for the post-survey.

Change in Frequency of Use of Standards-Based Instruction and Assessment.

Table 9 displays mean pre, post, and residual gain scores for teachers' use of standards-based instruction and assessment in mathematics. Over the course of the staff development initiative, teachers' mean use scores shifted from 44.39 to 46.65 on a scale of 14 (infrequent use of standards-based instruction and assessment) to 70 (frequent use of standards-based instruction and assessment). Furthermore, paired sample t-tests revealed that teachers' initial use of standards-based instruction and assessment was statistically significantly different from their post-use of instruction and assessment ($p < .001$). As a result, one can justifiably state that teachers reported use of standards-based strategies increased over the seven months in which this study was conducted. As in the previous discussion, the reader should note that the mean residual gain in level of use (1.39×10^{-15}) is not directly interpretable, as the mean of any set of residuals is always zero. As presented in the previous chapter, the coefficients of internal consistency for frequency of use of standards-based teaching strategies based on this sample were .90 for the pre-survey and .79 for the post-survey.

Table 9

Pre, Post, and Residual Gains in Use of Standards-Based Instruction and Assessment

	N	Minimum	Maximum	Mean	SD
Pre-Use of Standards-Based Instruction and Assessment	89	0	65	44.39	11.70
Post-Use of Standards-Based Instruction and Assessment	89	29	62	46.65	7.68
Residual Gain in Use of Standards-Based Instruction and Assessment	89	-17.52	16.55	1.39×10^{-15}	6.56
Valid N (listwise)	89				

Student surveys

Six hundred seven student surveys from 28 teachers' classes were completed. However, 6 of these 28 teachers did not complete their post-surveys, rendering it impossible to include their post-use of standards-based instruction and assessment scores in subsequent analyses of the correlation between teachers' reported use of standards-based instruction and assessment scores and students' perceptions of the use of standards-based instruction and assessment in their classrooms. Mean student scores of the 22 teachers' classes that were retained for analysis averaged 2.34 on a scale of 1 (frequent use of standards-based instruction and assessment in math) to 4 (infrequent use of standards-based instruction and assessment in math). These findings indicate that, on average, teachers are perceived by students as using standards-based teaching strategies

more than once a quarter, but not quite twice a month. The reliability of the scale used in the student survey was .59 for this sample.

Variables Excluded from Analysis

The independent variable, role in the staff development project, was excluded from this and subsequent analyses. It was the researcher's decision that deletion of this variable would increase the amount of information that potentially would be yielded in subsequent analyses. If, for example, role in the staff development project (lead versus participating teacher) was found to significantly predict residual gains in level of concern, intensity of concern, or frequency of use of standards-based teaching strategies, one would still be left with little information regarding the staff development activities, organizational contexts, or personal characteristics that were associated with changes in attitudes and practice. As a result, the major research questions of this study would remain unanswered. Deletion of the variable corresponding to teachers' roles also partially addressed the fact that lead teachers are represented in the sample at higher rates than in the population by reducing the emphasis in the study on the lead/participating teacher dichotomy.

Regression analysis assumes that independent variables are measured without error. As such, it is important to select the most reliable independent variables possible. As such, the internal reliability of each multi-item measure (the self-efficacy, principal directiveness, principal supportiveness, quality of school/community relations, organizational climate, and school-wide communication scales) was estimated with Cronbach's alpha. Reliability estimates of these scales are displayed in Table 10 below.

Table 10

Reliability Estimates of Self-Efficacy, Principal Directiveness, Principal Supportiveness, School/Community Relations, Organizational Climate, and School-Wide Communication Scales

Scale	Cronbach's α
Self-Efficacy Scale	.74
Principal Directiveness Scale	.62
Principal Supportiveness Scale	.86
School/Community Relations Scale	.48
Organizational Climate Scale	.59
School-Wide Communication Scale	.84

With a low reliability coefficient of .48, the School/Community Relations scale ($\alpha=.48$) was deemed inadequate for stepwise regression analysis. As a result, the quality of school/community relations variable was excluded from all subsequent analyses.

While the reliability estimates of the Organizational Climate and Principal Directiveness Scales (.59 and .62 respectively) are not that high, these variables were retained because more than half (approximately 60%) of the variance of these two variables was not due to error. Alpha reliability estimates in this range are generally considered adequate in survey research, particularly in cases such as this when the number of cases and items are relatively small.

Correlational analyses were employed to explore the relationships between the sixteen remaining independent variables and the three dependent variables. As shown in

Table 11, intercorrelations among the independent variables ranged from $-.19$ to $.79$. The low to moderate size of most correlation coefficients ($-.19$ to $.67$) suggests that the other independent variables were measuring different constructs. In addition, all intercorrelations among the dependent and independent variables range from $-.10$ to $.30$, suggesting that factors above and beyond the variables measured in this research are associated with gains in teachers' concerns and classroom practices.

Analysis of the Relationships among Independent Variables and Residual Gains in
Teachers' Level of Concern

Stepwise regression analyses were conducted to test the following null hypothesis:

Hypothesis 1: Changes in teachers' level of concern toward standards-based instruction and assessment in math cannot be predicted by

- a) teacher role in the staff development project (i.e., lead vs. participating teacher)
- b) size of the schools in which teachers work
- c) total number of years teaching experience
- d) teachers' sense of self-efficacy

Table 11

Zero Order Correlations among Criterion and Predictor Variables(Pairwise deletion of missing data, N=72-89)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Self-Efficacy	1.00																		
2. Years Teaching	.04	1.00																	
3. Standards-Based Classroom Activities	.15	-.06	1.00																
4. Action Research	-.14	.01	-.05	1.00															
5. Organiz. Climate	.10	-.16	-.09	-.03	1.00														
6. School-Wide Communic.	.17	-.06	.05	-.17	.67*	1.00													
7. Work with Consultant	-.16	-.10	-.16	-.03	.17	.03	1.00												
8. Curriculum Development	.05	-.12	-.16	-.08	.09	.08	.08	1.00											
9. Residual Gain in Intensity of Concern	.20	-.10	.09	.18	.00	.04	-.09	-.20	1.00										
10. Residual Gain in Use	.08	.30**	.28**	-.08	-.02	.19	-.13	-.19	.05	1.00									
11. Residual Gain in Level of Concern	.11	.02	.25*	.15	-.10	-.06	.09	.10	.10	.09	1.00								
12. Meetings	-.07	.12	-.17	.16	.10	.17	.28**	-.05	.17	-.04	-.06	1.00							
13. Principal Directiveness	-.16	.09	-.13	.32**	-.18	-.17	.03	.03	.05	.02	-.02	.04	1.00						
14. Principal Supportiveness	.17	-.07	.07	-.20	.59*	.79**	.01	.01	-.02	.20	-.10	.15	.23*	1.00					
15. Professional Reading	-.14	-.01	-.12	.05	-.11	-.13	.27*	-.03	-.03	.08	.04	-.03	.17	-.19	1.00				
16. Independent Research	.02	-.02	.06	-.02	.00	.09	.15	.20	.00	.12	.05	.22*	.02	.06	.04	1.00			
17. Size of School	.02	-.15	.27*	-.08	.16	.3**	-.15	-.07	-.05	.22*	-.07	.02	.17	.37**	-.07	.23*	1.00		
18. Examining Student Work	-.06	.02	-.06	.06	-.14	-.03	.11	.06	.05	.08	.22*	-.02	-.05	-.14	.06	.03	-.13	1.00	
19. Workshops	-.01	-.13	-.07	.11	.19	.07	.27*	.11	.12	.08	.15	.22*	-.01	.00	.17	.05	.06	-.06	1.00

Variables 9 through 11 are criterion variables.

*p≤.05 **p≤.01

- e) quality of school-wide communication
- f) degree of principal directiveness
- g) degree of principal supportiveness
- h) quality of school/community relations
- i) quality of organizational climate
- j) staff development activities in which teachers engage:
 - 1) action research
 - 2) workshops
 - 3) meetings/discussions with colleagues
 - 4) work with educational consultants
 - 5) carrying out standards-based activities in the classroom
 - 6) independent research
 - 7) professional reading
 - 8) curriculum development
 - 9) examining student work

In this analysis, the dependent variable, residual gain in level of concern, was regressed on the dependent variables listed in (b) through (g) and (i) through (j9) above. As stated earlier, the variables, (a) teacher role and (h) quality of school/community relations, were excluded from all analyses.

Assumptions of multiple regression were tested through a preliminary screening of the residuals. First, the studentized residuals were examined for any gross departures from normality. According to normal distribution theory, approximately 5% of outliers are expected to lie beyond +2 or –2 standard deviations. As shown in Table 12, 3/89 (or 3%) of the residuals in this analysis lie beyond these limits. As Table 12 shows, three individuals made gains that were much higher than predicted. However, the fact that fewer than 5% of outliers lie beyond ± 2 standard deviations places these findings well within the limits guiding the description of normality.

Table 12

Casewise Diagnostics (Dependent Variable=Residual Gain in Level of Concern)

Case Number	Studentized Residual	Residual Gain in Level of Concern	Predicted Value	Residual
53	2.21	2.59	-.48	3.07
55	3.05	3.68	-.55	4.22
60	2.92	3.68	-.37	4.05

Second, the normal probability plot was examined, which permits the researcher to compare the cumulative distribution of standardized residuals to the unit-normal distribution. If the data are from a normal distribution, the plotted values should fall roughly around the line. Non-linearity in the plot is evidence of non-normality.

Figure 1 displays the normal probability plot within two standard error limits. An examination of this plot reveals that the residual distribution is reasonably well-behaved.

The point at which residuals stray most from the fitted line is between .25 and .50 on the x-axis. In addition, it is evident that a small number of points fall outside the lower standard error limit. However, the cumulative distribution of standardized residuals does not differ enough from the unit-normal distribution to be considered non-normal.

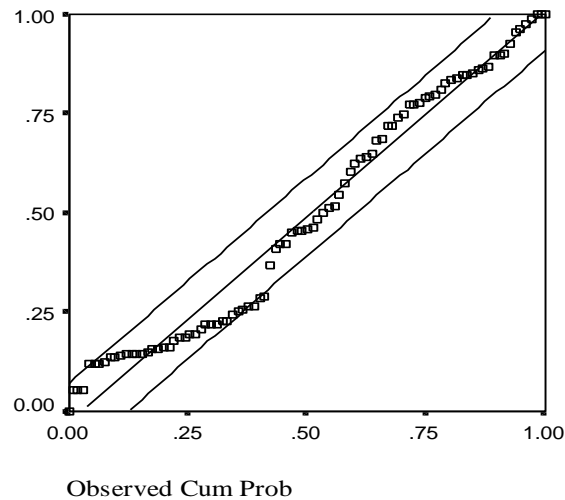


Figure 1. Normal probability plot of standardized residuals when independent variables are regressed on gain in level of concern

The assumption of homoscedasticity, or the assumption that standard deviations of errors of prediction are approximately equal for all predicted dependent variable scores, is a key assumption of multiple regression. To test the assumption of homoscedasticity, or equal variance, in regression analyses, it is recommended that the studentized residuals be plotted against the predicted values (Pedhazur, 1997; Glass & Hopkins, 1996; Tabachnik & Fidel, 1996), as shown in Figure 2. The plot in Figure 2

reveals that the residuals are not uniformly distributed. There is a cluster of points between -1 and approximately 0.5 on the x-axis and relatively few points beyond that point. As a result, the data appear to violate the assumption of homoscedasticity.

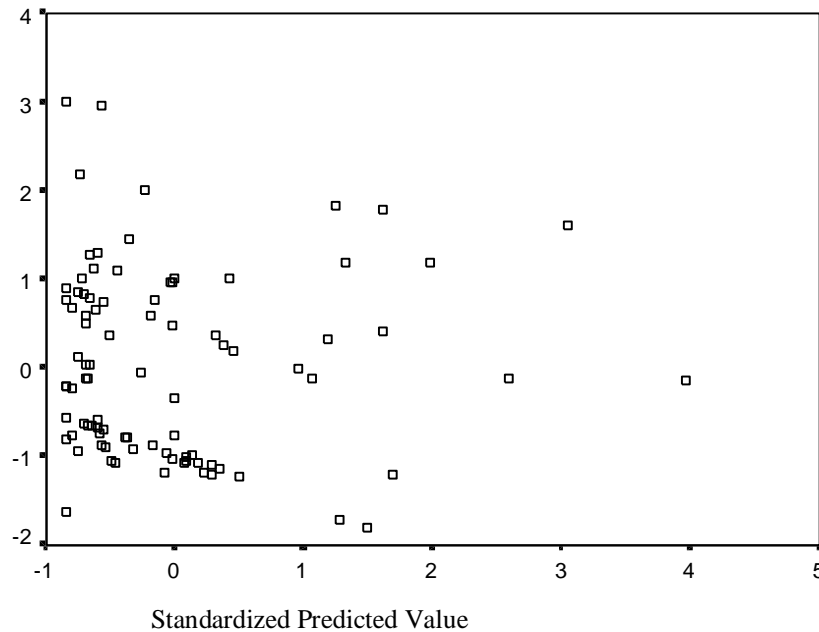


Figure 2. Plot of studentized residuals against standardized predicted values of gain in level of concern

The data were also examined to determine the degree to which they met the assumption that errors are not correlated. The fact that the Durbin Watson statistic of 2.34 was above the upper limit of 1.68 indicates that autocorrelation was not a problem in this data set (Chatterjee & Price, 1977, p. 223).

While not a formal assumption of multiple regression, ratio of cases to independent variables is a practical matter that must be addressed. Some researchers recommend that a case-to-independent variable ratio of 40-to-1 be utilized in stepwise

regression analyses (Tabachnik & Fidell, 1996). Obviously, the 4.6-to-1 ratio of case-to-independent variables in this analysis does not meet this recommendation. As stated previously, the risks assumed in a stepwise regression with a small sample are that the solution will be spurious, the results will not generalize beyond the sample, and/or that the analysis will lack adequate power to detect true differences.

In conclusion, this data set appears to meet all of the formal assumptions of multiple regression, except for one: homoscedasticity. However, it should be pointed out that heteroscedasticity does not invalidate a multiple regression analysis. Rather, heteroscedasticity merely threatens to weaken the analysis (Tabachnik & Fidell, 1996). As pointed out above, the real threat to this stepwise regression analysis is its small sample size.

With these points in mind, stepwise regression analysis was performed on this data set to explore the relative contribution of school size, teaching experience, self-efficacy, organizational climate, school-wide communication, principal directiveness, principal supportiveness, and 9 staff development activities in predicting teachers' residual gains in levels of concern about standards-based instruction and assessment. Results of this stepwise regression analysis revealed that a three variable model significantly predicts residual gains in level of concern ($F=4.9$, $p=.004$). As shown in Table 13, the strongest modifying influence on residual gains in level of concern were staff development activities: examining student work, implementing standards-based classroom activities, and carrying out action research. Eighteen percent of the variance in residual gains in concern was explained by a combination of these three variables.

Table 13

Summary of Stepwise Regression Analysis for Variables Predicting Residual Gains in Level of Concern (N=89)

Model	B	SE B	β	t	Sig.
1 .					
Examining Student Work	.102	.046	2.192	2.192	.032
2.					
Examining Student Work	.107	.045	.267	2.36	.021
Standards-Based					
Classroom Activities	.0114	.005	.240	2.13	.037
3.					
Examining Student Work	.103	.044	.256	2.32	.023
Standards-Based					
Classroom Activities	.012	.005	.245	2.22	.030
Action Research	.085	.040	.237	2.15	.035

Note. $R^2=.06$ for Model 1; $\Delta R^2=.06$ for Model 2; $\Delta R^2=.06$ for Model 3 ($p \leq .05$).

To reduce the risk that the statistical significance of this model ($F=4.9$, $p=.004$) was the result of a Type I error, Wilkinson and Dallal's tables of critical values for R^2 in forward stepwise selection were consulted. According Wilkinson and Dallal's tables (Tabachnik & Fidell, 1996, pp. 847-848), a multiple R^2 of approximately .18 is required to be significantly different from zero at $\alpha=.05$ in a forward stepwise regression in which there are 89 subjects, 16 potential independent variables, and an F-to-enter value of 3. Since the obtained R^2 of .18 equaled the critical R^2 , the hypothesis that this sample of scores was drawn from a population in which the multiple R was zero was rejected. Concerns that a Type I error had occurred were thus reduced.

Before interpreting the results displayed in Table 13, however, the issue of interactions was explored. Product terms were computed for the variables in the final regression equation presented in Table 13, representing the following two-way interactions: student work/ action research, student work/standards-based classroom activities, and action research/ standards-based classroom activities.

A direct entry multiple regression analysis was then conducted, regressing residual gains in level of concern on student work, standards-based classroom activities, action research, and the three interaction terms listed above. The results of this analysis are presented in Table 14.

The resulting regression equation was:

$$\begin{aligned} \text{Residual gains in level of concern} = & -.37 + .01 * \text{standards-based classroom activities} - .07 * \text{action} \\ & \text{research} + .108 * \text{student work} - .005 * \text{student work/standards-} \\ & \text{based classroom activities} + .007 * \text{student work/action} \\ & \text{research} + .008 * \text{action research/ standards-based classroom} \\ & \text{activities} \end{aligned}$$

The estimated standard errors for b_1 through b_6 were .005, .069, .054, .004, .017, and .004, respectively. T-Tests of the b_1 , b_2 , and b_3 coefficients revealed that only the standards-based classroom activities ($t=2.21$, $p=.03$) and examining student work ($t=2.01$, $p=.048$) variables contributed significantly to prediction. Further t-tests of the b_4 , b_5 , and b_6 coefficients yielded a statistically significant result for b_6 ($t=2.3$, $p=.024$). This suggested the presence of an interaction effect between action research and the implementation of standards-based classroom activities.

Table 14

Summary of Stepwise Regression Analysis for Independent Variables and Interaction Variables Predicting Residual Gains in Level of Concern (N=89)

Model	B	SE B	β	t	Sig.
1.					
Standards-Based Classroom Activities	.01	.005	.23	2.21	.03
Action Research	-.07	.069	-.19	-1.01	.27
Examining Student Work	.11	.054	.24	2.01	.048
Examining Student Work/Standards-Based Classroom Activities Interaction	-.005	.004	-.28	-1.29	.20
Examining Student Work/Action Research Interaction	.007	.017	.08	.40	.691
Action Research/Standards-Based Classroom Activities Interaction	.008	.004	.50	2.30	.024
<u>Note:</u> $R^2=.198$ for Model 1.					

The strength of the effect was indexed by the difference in squared multiple correlations for the “main effects only” model and the interaction model. For the former, the squared multiple correlation was .178, and for the latter it was .198. This yielded $.198-.178=.02$. The interaction effect therefore accounted for 2% of the variance in residual gains in level of concern, a relatively small effect size.

To examine the nature of the interaction, the implementation of standards-based classroom activities was used as the moderator variable. With no theory to guide the selection of a moderator variable, this decision was made in an attempt to study more closely the effects of action research, a staff development activity in which just 40% of teachers participated. Since 84% of all teachers reported that they had spent time implementing standards-based activities in their classrooms, the researcher felt that it would be more interesting to examine how the relationship between residual gains in level of concern and action research varied across hours spent implementing standards-based classroom activities. The value of b_6 therefore indicated this relationship. For every hour that action research was implemented, the slope of residual gains in level of concern on standards-based classroom activities changed by .008 units.

To investigate the nature of this interaction more thoroughly, “low” and “high” scores for the standards-based classroom activities variable were calculated. Using the values in Table 5, a “low” score was defined as one standard deviation (31.01 hours) below the mean standards-based activities score (17.22 hours). A “high” score was similarly defined as one standard deviation above the mean standards-based activities score. Thus, the values of low and high scores on the standards-based classroom activity variable were -13.79 and 48.23 , respectively². The slope of residual gains in level of concern on action research when implementation of standards-based classroom activity was “average” (i.e., at the mean) was obtained from the above regression equation,

² The value of a low score on the standards-based classroom activity variable (-13.79 hours) is theoretical, as it is impossible for teachers to spend fewer than zero hours engaged in this activity. Nevertheless, this value was utilized to explore the nature of the action research/standards-based activities interaction at the

namely b_2 , which equaled $-.07$. The slopes of residual gains in level of concern on action research for the cases of “low” and “high” implementation of standards-based classroom activities were calculated, using the following equation for calculating the slope of the predicted effects of action research (X_2) on residual gains in level of concern (Y) at any particular value of implementation of standards-based classroom activities (X_1), where b_2 =the regression coefficient for action research and b_6 =the regression coefficient for the standards-based activities/action research interaction.

$$b_2 \text{ at } X_1 = b_2 + b_6 X_1 \quad (\text{Jaccard, Turrisi, \& Wan, 1990, p. 26})$$

$$b_2 \text{ at } -13.79 = -.07 + (.008)(-13.79) = -.18$$

$$b_2 \text{ at } 48.23 = -.07 + (.008)(48.23) = .32$$

As presented in Table 14, the standard error for b_2 when implementation of standards-based activities was average was $.069$. The standard errors for b_2 at “low” and “high” values of implementing standards-based classroom activities were calculated as recommended in Jaccard, Turrisi, & Wan (1990, p. 27):

$$\begin{aligned} s(b_2 \text{ at } -13.79) &= [\text{var}(b_2) + (-13.79)^2 \text{var}(b_6) + 2(-13.79)\text{cov}(b_2, b_6)]^{1/2} \\ &= [.005 + (190.16)(.00001) + (27.58)(-.00002)]^{1/2} \\ &= [.005 + .0019 - .005516]^{1/2} \\ &= [.001384]^{1/2} \\ &= .037 \end{aligned}$$

three levels of the moderator variable that are routinely investigated in interaction analysis: the mean value, one standard-deviation below the mean, and one standard-deviation above the mean.

$$\begin{aligned}
s(b_2 \text{ at } 48.23) &= [\text{var}(b_2) + (48.23)^2 \text{var}(b_6) + 2(48.23)\text{cov}(b_2, b_6)]^{1/2} \\
&= [.005 + (2326.13)(.00001) + (94.46)(-.0002)]^{1/2} \\
&= [.005 + .023 - .018892]^{1/2} \\
&= [.009108]^{1/2} \\
&= .095
\end{aligned}$$

The t ratios for the “low,” “average,” and “high” standards-based activities values were computed by dividing the regression coefficients by their standard errors. The results of these calculations can be summarized as follows in Table 15:

Table 15

Regression Coefficients, Standard Errors, and T-Values for Action Research at High, Average, and Low Rates of Implementation of Standards-Based Activities in the Classroom

Implementation Rates	b_2	SE	t
Low Implementation	-.18	.037	-4.86**
Average Implementation	-.07	.07	2.2
High Implementation	.32	.095	3.36**

** p<.01

Using Bonferroni corrected alpha levels, the t values for b_2 at low and high levels of implementation of standards-based activities were found to be statistically significant at the .01 level (barring a Type I error). As demonstrated previously in Table 14, the regression coefficient for action research at an average level of implementation of standards-based classroom activities was not found to differ significantly from zero.

These results were subsequently interpreted as follows: At “average” levels of implementation of standards-based classroom activities, the effect of each additional hour of action research was insignificant. At “low” levels of implementation of standards-based classroom activities, however, each additional hour of action research was associated with a .18 decrease in residual gains in level of concern. Furthermore, at “high” levels of implementation of standards-based activities in the classroom, each additional hour of action research was associated with an increase in residual gains in level of concern of .32. These findings suggest that action research might play a beneficial role (in terms of residual gain in level of concern) when teachers are heavily involved in the implementation of standards-based classroom activities. On the other hand, teachers who are not implementing these activities in the classroom yet undertake an action research project are likely to experience negative residual gains in level of concern.

Although the main effect for the implementation of standards-based activities was statistically significant (see Table 14), this main effect was not interpreted because the interaction between action research and the implementation of standards-based activities in the classroom was disordinal (i.e., when each variable was plotted against residual gain in level of concern, the regression line for action research intersected the regression line for standards-based activities). As a result, the main effects of implementing standards-based activities in the classroom were too confounded with the effects of action research to be interpreted separately.

Figure 3 displays overlapping regression lines for action research and the implementation of standards-based classroom activities, when residual gains in level of concern are regressed (separately) on each variable. The plot in Figure 3 represents a disordinal interaction between action research and implementing standards-based activities.

Finally, the main effects of examining student work were addressed. According to the model displayed in Table 14, one can postulate that (at least for this data set) for every one-hour increase in activities involving the examination of student work among small groups of teachers, residual gains in level of concern increase by .108 (holding the effect of the interaction between classroom activities and action research variables constant). From these results, it is possible to predict that teachers will move from one level of concern to a higher level of concern (e.g. from level 1 to level 2) upon the completion of approximately 9.3 hours examining student work. This could provide useful information to staff developers and change agents who are planning an innovation, for these results indicate that the small group activities focusing on examining student work are associated with relatively “quick” residual gains in level of concern..

It is interesting that only staff development activities were found to significantly predict residual gains in level of concern. No personal or contextual variables were significant predictors, suggesting, perhaps, that staff development alone is sufficient to bring about changes in level of concern.

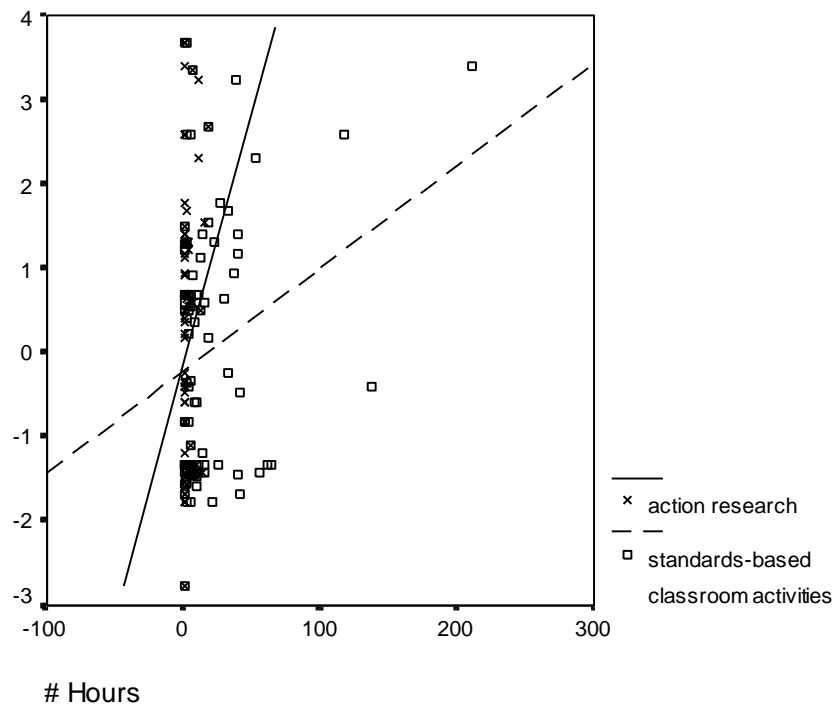


Figure 3. Disordinal interaction between the action research and standards-based classroom activities variables

Although these regression results indicate that 19.8% of the variability in residual gains in level of concern may be predicted by the degree to which teachers examine student work, and the interaction between implementing standards-based activities in the classroom and participating in action research projects, almost 80% of the variance in residual gains in level of concern is still unaccounted for. As such, it appears that some key variables have been omitted from this study. This topic merits further research.

Analysis of the Relationships among the Independent Variables and Residual Gains an Teachers' Intensity of Concern

A stepwise regression was conducted to test the following null hypothesis:

Hypothesis 2: Changes in teachers' intensity of concern toward standards-based instruction and assessment in math cannot be predicted by

- a) teacher role in the staff development project (i.e., lead vs. participating teacher)
- b) size of the schools in which teachers work
- c) total number of years teaching experience
- d) teachers' sense of self-efficacy
- e) quality of school-wide communication
- f) degree of principal directiveness
- g) degree of principal supportiveness
- h) quality of school/community relations
- i) quality of organizational climate
- j) staff development activities in which teachers engage:
 - 1) action research
 - 2) workshops
 - 3) meetings/discussions with colleagues
 - 4) work with educational consultants
 - 5) carrying out standards-based activities in the classroom
 - 6) independent research

- 7) professional reading
- 8) curriculum development
- 9) examining student work

In this analysis, the dependent variable, residual gain in intensity of concern, was regressed on the independent variables listed in (b) through (g) and (i) through (j9) above. As stated earlier, the variables, (a) teacher role and (h) quality of school/ community relations, were excluded from all analyses.

As in the previous analysis, assumptions of multiple regression were tested through a preliminary screening of the residuals. First, the studentized residuals were examined for any gross departures from normality. As displayed in Table 16, 5/89 (or 5.6%) of the residuals in this analysis lie beyond these limits. These studentized residuals range from -2.5 to 2.37 . As Table 18 shows, one individual made gains that were much higher than predicted, while four others made much smaller gains than expected. Under the assumption of normality, between 4 and 5 cases with studentized residuals (2 standard deviations would not be considered unusual (5% of 89 is 4.5 cases). With this in mind, the presence of 5 cases with studentized residuals greater than ± 2 was not considered to violate the assumption of normality.

Secondly, the normal probability plot was examined, to compare the cumulative distribution of standardized residuals to the unit-normal distribution. Figure 4 displays the normal probability plot within two standard error limits. An examination of this plot reveals that the residual distribution is very well-behaved. The points cling very

closely to the regression line, and very few are observed beyond the two standard error limit.

Table 16

Casewise Diagnostics (Dependent Variable=Residual Gain in Intensity of Concern)

Case Number	Studentized Residual	Residual Gain in Intensity of Concern	Predicted Value	Residual
4	2.37	34.46	-6.22	40.68
18	-2.28	-34.04	5.12	-39.16
54	-2.36	-34.42	6.08	-40.50
56	-2.50	-66.04	-23.03	-43.01
84	2.42	36.77	-4.77	41.54

To test the assumption of homoscedasticity, or equal variance, the studentized residuals were plotted against the predicted values, as shown in Figure 5. The plot in Figure 5 reveals no problem with heteroscedasticity. Except for three points between – 3.5 and –2 on the x-axis, one observes the characteristic “blob” expected when residuals are fairly uniformly distributed.

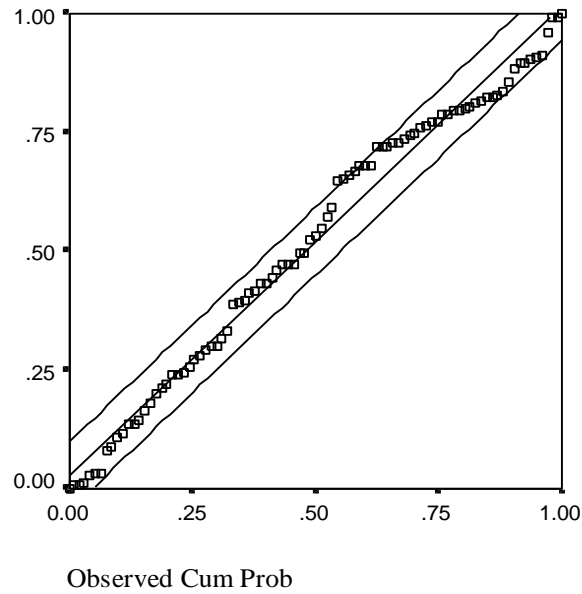


Figure 4. Normal probability plot of standardized residuals when independent variables are regressed on gain in intensity of concern

The data were also examined to determine the degree to which they met the assumption that errors are not correlated. The fact that the Durbin Watson statistic of 1.84 was

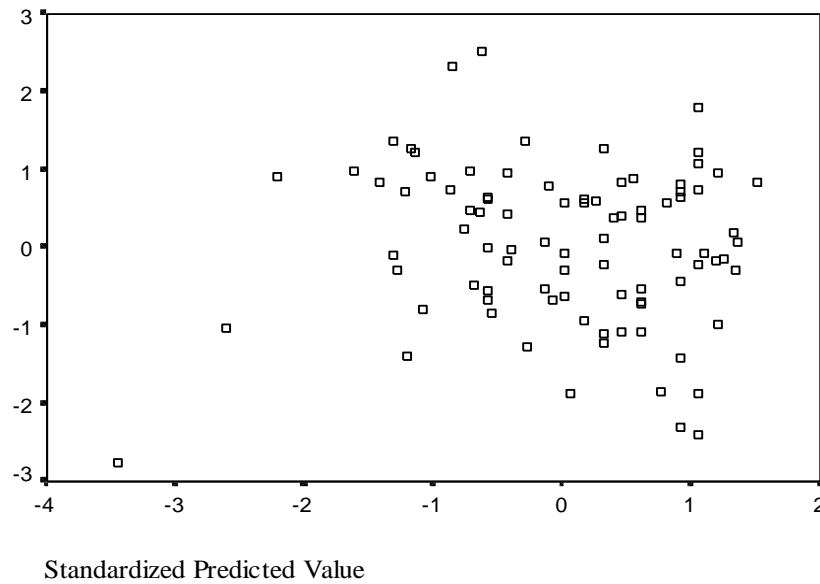


Figure 5. Plot of studentized residuals against standardized predicted values for gain in intensity of concern

above the upper limit of 1.68 indicates that autocorrelation was not a problem in this data set (Chatterjee & Price, 1977, p. 223).

In conclusion, this data set appears to meet all of the formal assumptions of multiple regression. However, all of the limitations and risks associated with conducting stepwise regression analyses that were mentioned in the previous section hold true for this analysis, as well.

Stepwise regression analysis was subsequently performed to examine the relative contribution of school size, teaching experience, self-efficacy, organizational climate, school-wide communication, principal directiveness, principal supportiveness, and 9 staff development activities in predicting teachers' residual gains in intensity of concern about standards-based instruction and assessment. Results of this stepwise regression analysis revealed that a two variable model significantly predicted residual gains in level of concern ($F=4.996$, $p=.009$). As shown in Table 17, the strongest modifying influence on residual gains in intensity of concern was teacher self-efficacy, followed by involvement in curriculum development. Almost thirteen percent of the variance in residual gains in intensity of concern was explained by a combination of these two variables.

As in the first analysis, Wilkinson and Dallal's tables of critical values for R^2 in forward stepwise selection were consulted, to reduce the risk that the statistical significance of this model had occurred as a result of a Type I error. Since a multiple R^2 of approximately .18 is required to be significantly different from zero at $\alpha=.05$ in a

forward stepwise regression in which there are 89 subjects, 16 potential independent variables, and an F-to-enter value of 3, it was obvious that the obtained R^2 of .13 did not

Table 17

Summary of Stepwise Regression Analysis for Variables Predicting Residual Gains in Intensity of Concern (N=89)

Model	B	SE B	β	t	Sig.
1. Self-Efficacy	14.92	6.40	.27	2.33	.023
2. Self-Efficacy	15.33	6.26	.28	2.45	.017
Curriculum Development	-1.31	.63	-.23	-2.07	.042

Note. R^2 =.072 for Model 1; ΔR^2 =.054 for Model 2 (p s<.05).

meet this criteria. Hence, it was impossible to conclude (at the .05 level) that an R^2 of .13 was actually different from zero. However, the obtained R^2 of .13 was larger than the critical R^2 of approximately .12 that is required for R^2 to be statistically significant at the .10 level (when $N=70$, $k=16$, and F -to-enter =3).

Obviously, these findings are not ideal. In most social science research, it is customary to report statistically significant results only when the researcher is 95% to 99% certain that his/her findings are not due to sampling error. Here, the statistical significance of the model yielded to predict residual gains in intensity of concern runs a 10% risk of occurring due to error. Despite this fact, a decision was made to proceed with this analysis. This study is purely exploratory, no previous studies have attempted to predict changes in attitude from staff development, contextual, and personal variables, and any ideas gleaned from this analysis will certainly need to be put to the test in future studies. Similarly, Licht (1997) concluded:

“In general, ‘fishing expeditions,’ in which variables are included because they might be useful, are discouraged because they are likely to result in highly inflated Type I error rates; although, in preliminary stages of the study of a phenomenon, these types of exploratory investigations can prove useful. When they are used, however, they should be clearly labeled as exploratory, statistical significance should be interpreted with extreme caution, and results should be replicated in more carefully designed confirmatory studies” (p. 55).

As a result, the decision was made to explore this model, notwithstanding the fact that the risk of a Type I error was slightly higher in this analysis than for the previous one involving residual gains in level of concern.

At this point, the issue of interactions was explored. One product term was computed for the two variables in the final regression equation presented in Table 17, representing the self-efficacy/curriculum development interaction. A direct entry multiple regression analysis was then conducted regressing residual gains in intensity of concern on self-efficacy, curriculum development, and the self-efficacy/curriculum development interaction. The resulting regression equation was:

$$\text{Residual gain in intensity of concern} = -32.63 + 9.77 * \text{self-efficacy} - 13.293 * \text{curriculum development} + 3.346 * \text{self-efficacy/curriculum development}$$

The estimated standard errors for b1, b2, and b3 were 6.80, 6.30, and 1.75, respectively. A test of the b3 coefficient did not yield a statistically significant result (t=1.91, p=.06). This suggests a lack of an interaction effect for this model. With no significant interaction effects, interpretation of the main effects displayed in Table 17 was straightforward. The regression equation for this model was:

$$\text{Residual gain in intensity of concern} = -52.37 + 15.33 * \text{self-efficacy} - 1.31 * \text{curriculum development}$$

According to the model displayed in Table 17, it appears that for every single unit increase in self-efficacy (on a scale of 1 (low self-efficacy) to 4 (high self-efficacy)), residual gains in intensity of concern increase by 15.33 (holding the effect of curriculum development constant). Similarly, residual gains in intensity of concern actually decrease by 1.31 for every hour spent developing curriculum (holding the effect of self-efficacy constant).

As mentioned previously, intensity of concern is not in and of itself directly interpretable. For example, high intensity of concern could indicate interest, curiosity, or enthusiasm about standards-based instruction and assessment. By the same token, high intensity of concern could correspond with strong feelings of resentment, stress, or anger about this innovation. If intensity of concern is desired (i.e., it denotes enthusiasm, commitment, or other positive attitudes), these results suggest that staff development aimed at increasing teachers' self-efficacy might contribute to increased intensity of concern about standards-based instruction and assessment in math. Even if self-efficacy training or therapy led to more intense, yet negative, concerns about standards-based instruction and assessment in mathematics, it would be hard to argue that teachers' self-efficacy should be kept low.

An interesting feature of this model is that less involvement in curriculum development is associated with gains in intensity of concern, a fact that was corroborated by information volunteered by teachers in teacher interviews. Teachers in the sample

who did not develop or modify standards-based math curriculum were, on the whole, more positive about standards-based reform in math. Those who attempted to develop curriculum were frustrated by the difficulty of this exercise and, in turn, became less enthusiastic about the future of standards-based instruction and assessment in mathematics. Based on teacher interviews, it became clear that, at least for this sample of teachers, less curriculum development was a good thing.

As in the previous analysis, no contextual variables (i.e., quality of organizational climate, quality of school-wide communication, degree of principal directiveness, or degree of principal supportiveness) significantly predicted residual gains in intensity of concern. These findings, in combination with the previous results of the regression analysis on residual gains in level of concern, suggest that overall changes in concern might be brought about through staff development activities and personal factors, regardless of the context in which teachers work.

It must be kept in mind that these regression results indicate that just 13% of the variability in residual gains in intensity of concern may be predicted by the teachers' self-efficacy and their involvement in curriculum development. As in the previous analysis, a substantial portion of the variance in residual gains in intensity of concern (87%) failed to be accounted for in this model, signifying that other factors play a larger role in producing residual gains in intensity of concern. Finally, the reader must be aware that the risk of a Type I error for this analysis is 10%, which is higher than in the first analysis involving residual gains in level of concern.

Analysis of the Relationships among the Independent Variables and Residual Gains in
Teachers' Reported Use

A stepwise regression was conducted to test the following null hypothesis:

Hypothesis 3: Changes in teachers' reported use of standards-based instruction and assessment in math cannot be predicted by

- a) teacher role in the staff development project (i.e., lead vs. participating teacher)
- b) size of the schools in which teachers work
- c) total number of years teaching experience
- d) teachers' sense of self-efficacy
- e) quality of school-wide communication
- f) degree of principal directiveness
- g) degree of principal supportiveness
- h) quality of school/community relations
- i) quality of organizational climate
- j) staff development activities in which teachers engage:
 - 1) action research
 - 2) workshops
 - 3) meetings/discussions with colleagues
 - 4) work with educational consultants

- 5) carrying out standards-based activities in the classroom
- 6) independent research
- 7) professional reading
- 8) curriculum development
- 9) examining student work

In this analysis, the dependent variable, residual gain in use of standards-based instruction and assessment in math, was regressed on the independent variables listed in (b) through (g) and (i) through (j9) above. As stated earlier, the variables, (a) teacher role and (h) quality of school/community relations, were excluded from all analyses.

To investigate the concurrent validity of teachers' reported use of standards-based instruction and assessment, correlations were run between teachers' raw scores on the post-version of the use of standards-based instruction and assessment scale and the mean student-reported score of the same construct. This analysis revealed a statistically significant correlation ($r = -.56$, $p = .006$) between teachers' reported use of standards-based instruction and assessment and students' perceptions of what goes on in their math classes³. These findings lent considerable validity to the use of teachers' raw scores on the post-version of the use of standards-based instruction and assessment scale as a proxy for teachers' actual use of standards-based instruction and assessment in their math classes.

³ The negative correlation is due to the direction of the scales used in the teacher and student surveys. On the teacher survey, a high score indicated frequent use of standards-based instruction and assessment. On the student survey, a low score indicated student perception of frequent use of standards-based instruction and assessment.

In the next analysis, the dependent variable, residual gain in use of standards-based instruction and assessment in math, was regressed on the independent variables listed in (b) through (g) and (i) through (j10) above. As stated earlier, the variables, (a) teacher role and (h) quality of school/community relations, were excluded from all analyses.

As in the previous analyses, assumptions of multiple regression were tested through a preliminary screening of the residuals. Studentized residuals were examined for any gross departures from normality. As shown in Table 18, 2/70 (or 2.9%) of the residuals in this analysis lie beyond these limits. The existence of 2.9% of cases with studentized residuals greater than ± 2 is well within the guidelines of normality.

Table 18

Casewise Diagnostics (Dependent Variable=Residual Gain in Reported Use of Standards-Based Instruction and Assessment)

Case Number	Studentized Residual	Residual Gain in Reported Use	Predicted Value	Residual
56	-2.14	-17.52	-5.84	-12.03
87	-2.67	-11.75	3.23	-14.98

Next, the normal probability plot was examined, to compare the cumulative distribution of standardized residuals to the unit-normal distribution. Figure 6 displays the normal probability plot within two standard error limits. An examination of this plot reveals that the residual distribution is extremely well-behaved, as the points cling very closely to the regression line. This fact further suggests evidence of normality.

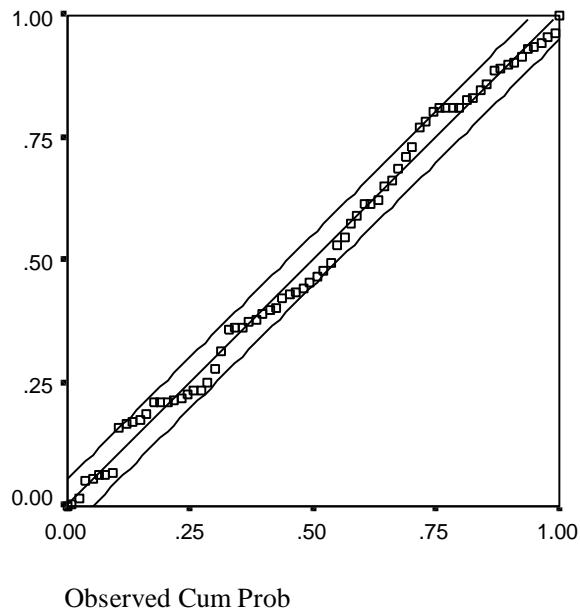


Figure 6. Normal probability plot of standardized residuals when independent variables are regressed on gain in level of use

To test the assumption of homoscedasticity, the studentized residuals were plotted against the predicted values, as shown in Figure 7.

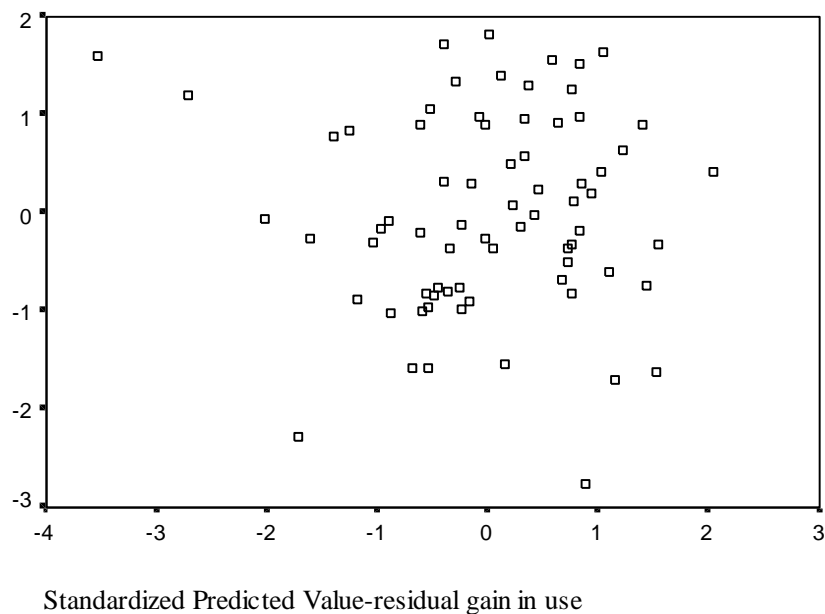


Figure 7. Plot of studentized residuals against predicted values of residual gain in use

The plot in Figure 7 reveals no problem with heteroscedasticity. Except for three points between -3.5 and -2 on the x-axis, the overall shape of the scatterplot does not violate the assumption of homoscedasticity.

The data were also examined to determine the degree to which they met the assumption that errors are not correlated. The fact that the Durbin Watson statistic of 1.84 was above the upper limit of 1.745 indicates that autocorrelation was not a problem in this data set (Chatterjee & Price, 1977, p. 223).

In conclusion, this data set appears to meet all of the formal assumptions of multiple regression. As with the previous analyses, the greatest threat to the interpretability of subsequent stepwise regression results is the extremely small sample size.

Stepwise regression analysis subsequently was performed to determine the relative contribution of school size, teaching experience, self-efficacy, organizational climate, school-wide communication, principal directiveness, principal supportiveness, and 9 staff development activities in predicting teachers' residual gains in use of standards-based instruction and assessment. Results of this stepwise regression analysis revealed that a four variable model significantly predicts residual gains in use of standards-based instruction and assessment ($F=6.35$, $p<.001$). As shown in Table 19, the strongest modifying influences on residual gains in use of standards-based instruction and assessment was years teaching, followed by degree of principal supportiveness, quality of organizational climate, and involvement in curriculum development. Almost

twenty-eight percent of the variance in residual gains in use was explained by a combination of these four variables.

As in the previous analyses, Wilkinson and Dallal's tables of critical values for R^2 in forward stepwise selection were consulted. According to Wilkinson and Dallal's tables (Tabachnik & Fidell, 1996, pp. 847-848), a multiple R^2 of approximately .22 is required to be significantly different from zero at $\alpha=.05$ in a forward stepwise regression in which there are 70 subjects, 16 potential independent variables, and an F-to-enter value of 3. Since the obtained R^2 of .28 exceeded the critical R^2 , the hypothesis that this sample of scores was drawn from a population in which the multiple R was zero was rejected. Concerns that a Type I error had occurred were thus reduced.

Prior to interpreting the results displayed in Table 19, interactions were explored. Product terms were computed for the variables in the final regression equation presented in Table 19, representing the following two-way interactions: years teaching/principal supportiveness, years teaching/organizational climate, years teaching/curriculum development, principal supportiveness/organizational climate, principal supportiveness/curriculum development, and organizational climate/curriculum development.

A direct entry multiple regression analysis was then conducted regressing residual gains in use on years teaching, principal supportiveness, organizational climate, curriculum development, and the six interaction terms listed above. The resulting regression equation was:

Table 19

Summary of Stepwise Regression Analysis for Variables Predicting Residual Gains in Use of Standards-Based Instruction and Assessment (N=70)

Model	B	SE B	β	t	Sig.
1. Years Teaching	1.38	.52	.304	2.669	.009
2. Years Teaching	1.45	.508	.319	2.858	.006
Standards-Based Classroom Activities	.047	.023	.232	2.078	.041
3. Years Teaching	1.52	.499	.333	3.04	.003
Standards-Based Classroom Activities	.045	.022	.220	2.011	.048
Principal Supportiveness	2.34	1.20	.213	1.944	.056
4. Years Teaching	1.32	.49	.290	2.698	.009
Standards-Based Classroom Activities	.034	.022	.167	1.545	.127
Principal Supportiveness	4.50	1.48	.410	3.034	.003
Organizational Climate	-4.96	2.11	-.324	-2.353	.022
5. Years Teaching	1.20	.49	.263	2.475	.016
Standards-Based Classroom Activities	.028	.022	.138	1.285	.203
Principal Supportiveness	4.67	1.46	.425	3.206	.002
Organizational Climate	-5.07	2.07	-.331	-2.454	.017
Curriculum Development	-.41	.21	-.204	-1.932	.058
6. Years Teaching	1.13	.484	.248	2.337	.022
Principal Supportiveness	5.00	1.44	.455	3.47	.001
Organizational Climate	-5.63	2.03	-.367	-2.77	.007
Curriculum Development	-.44	.209	-.223	-2.127	.037

Note. $R^2=.092$ for Model 1; $\Delta R^2=.053$ for Model 2; $\Delta R^2=.045$ for Model 3; $\Delta R^2=.062$ for Model 4; $\Delta R^2=.040$ for Model 5 ($p<.05$); $\Delta R^2=-.018$ for Model 6 ($p=.2$).

$$\begin{aligned} \text{Residual gains in use} = & -12.46 + 5.86* \text{ principal supportiveness} + 1.45* \text{ organizational climate} \\ & + 1.32* \text{ curriculum development} + .41* \text{ years teaching} + .83* \text{ years} \\ & \text{teaching/principal supportiveness} -.52* \text{ years teaching/organizational} \\ & \text{climate} -.22* \text{ years teaching/curriculum development} -1.56* \text{ principal} \\ & \text{supportiveness/organizational climate} +.51* \text{ principal} \\ & \text{supportiveness/curriculum development} -.96 * \text{ organizational} \\ & \text{climate/curriculum development} \end{aligned}$$

Regression coefficients, standard errors, t-values, and significance test results for this model are displayed in Table 20. As displayed in Table 20, tests of the regression coefficients for the interaction terms did not yield statistically significant results. This suggests a lack of an interaction effect for this model.

With no significant interaction effects, interpretation of the main effects displayed in Table 19 was straightforward. The regression equation for this model was:

$$\begin{aligned} \text{Residual gain in use} = & -2.32 + 5.00* \text{ principal supportiveness} - 5.63* \text{ organizational climate} \\ & - .44 * \text{ curriculum development} + 1.13* \text{ years teaching} \end{aligned}$$

Common sense says that it is more difficult to change the way one behaves than it is simply to change one's attitudes. A person can want to lose weight but find it hard to stop eating fattening food. Another individual may strive for an orderly house but find it nearly impossible to pick up after him/herself. By the same token, previous research has shown that it is easier for teachers to change their attitudes than it is for them to change their everyday teaching strategies. This is validated by the fact that prediction of

Table 20

Regression coefficients, standard errors, t-values, and significant test results for the interaction model (N=70)

Interaction Model	B	SE B	t	Sig.
Years Teaching	5.86	10.35	.57	.57
Principal Supportiveness	1.45	11.67	.12	.90
Organizational Climate	1.32	3.24	.41	.69
Curriculum Development	.41	3.68	.11	.91
Years Teaching/Principal Supportiveness	.83	1.02	.82	.42
Years Teaching/Organizational Climate	-.52	1.57	-.33	.74
Years Teaching/Curriculum Development	-.22	.19	-1.14	.26
Principal Supportiveness/Organizational Climate	-1.55	3.40	-.46	.65
Principal Supportiveness/Curriculum Development	.51	.80	.63	.53
Organizational Climate/Curriculum Development	-.96	1.32	-.73	.47

residual gains in frequency of use of standards-based instruction and assessment is more complex than it is for prediction of residual gains in level and intensity of concern. The

results of this analysis revealed that a combination of personal (years teaching), organizational (principal supportiveness and organizational climate), and staff development (curriculum development) factors were influential in predicting residual gains in use of standards-based instruction and assessment.

According to the model displayed in Table 19, it is possible to conjecture that (at least for this data set) for every single unit increase on the 5-point teaching experience scale (1=less than 1 year; 2=2-5 years; 3=6-10 years; 4=11-19 years; 5=20 or more years), residual gains in use increase by 1.13 (holding the effect of the other variables in the model constant) on the 56-point frequency of use scale. Similarly, residual gains in use increase by 5.00 for every single unit increase in principal supportiveness (holding the other variables constant). This would correspond to an approximate 8.9% increase in the use of standards-based practices in the classroom. Furthermore, residual gains in use seem to decrease by 5.63 for each additional unit increase in quality of organizational climate (holding the effect of the other variables constant). Finally, residual gains in use fall by .44 for every hour of curriculum development conducted.

Perhaps the most curious feature of these results is the fact that residual gains in use actually decline as the quality of a school's organizational climate improves. It is also interesting that organizational climate was a significant predictor of residual gain in use at all since it is virtually uncorrelated with residual gain in use ($r=-.02$). It is possible, therefore, that organizational climate played the role of suppressor variable in this analysis. According to Horst (1941), a suppressor variable is a variable that has a zero, or close to zero correlation with the criterion yet leads to improvement in prediction when

included in a multiple regression analysis. At the same time, a suppressor variable is correlated with one or more of the other predictor variables. A suppressor variable works by controlling for, or suppressing, irrelevant variance (i.e., variance it shares with the predictors and not with the criterion), thereby ridding the analysis of irrelevant variation.

Horst (1966) also explained:

“A suppressor variable may be defined as those predictor variables which do not measure variance in the criterion measures, but which do measure some of the variance in the predictor measures which is not found in the criterion measure. They measure invalid variance in the predictor measures and serve to suppress this invalid variance (p. 363).”

As stated above, the organizational climate variable is almost completely uncorrelated with the criterion for this analysis, residual gain in frequency of use. In addition, it is highly correlated with the degree of principal supportiveness variable ($r=.59$, $p<.05$). It displays very low correlations with the years teaching and curriculum development variables ($r=-.16$ and $.09$, respectively), the other significant predictors. Consequently, quality of organizational climate is related to principal supportiveness, and whatever these two variables share is different from what the organizational climate variable shares with the years teaching curriculum development variables. In this analysis, organizational climate likely served as a suppressor variable by suppressing irrelevant variance in the principal supportiveness variable and consequently enhancing prediction of residual gain in use by the principal supportiveness variable.

According to Pedhazur (1997), it is typical for a suppressor variable to have a negative regression coefficient, as was the case with organizational climate ($b=-5.63$).

He explains:

“Consequently, when the regression equation is applied, predicted scores for people who score above the mean on the suppressor variable are lowered as a result of multiplying a negative regression coefficient by a positive [standard] score. Conversely, predicted scores of those below the mean on the suppressor variable are raised as a result of multiplying a negative regression coefficient by a negative [standard] score. In other words, people who are high on the suppressor variable are penalized, so to speak, for being high, whereas those who are low on the suppressor variable are compensated for being low (p. 187).”

This explains the rather counterintuitive results yielded in this analysis, namely that increases in organizational climate are associated with decreases in use of standards-based teaching and assessment practices. While the organizational climate variable’s curious function in the prediction of residual gains in frequency of use is therefore more clear, its role as a suppressor variable prevents it from being one that is meaningfully interpreted on its own. Rather, its use lies in its ability to enhance the prediction of the principal supportiveness variable.

Except for the curriculum development, the variables in the regression equation predicting residual gains in frequency of use are rather difficult to manipulate. For example, principal supportiveness is not a factor over which teachers have much control. While these results suggest that increases in the degree of principal supportiveness a teacher experiences are associated with gains in use of standards-based teaching strategies, this independent variable is not easy to change and requires a significant amount of time and effort on the part of change agents.

A final finding of this analysis is that teachers’ residual gains in use increase by 1.13 for every unit increase on the teaching experience scale. This suggests that older teachers in the study sample changed their mathematics teaching strategies more than younger, less experienced teachers. Using the findings in this study, one may predict that

residual gains in frequency of use of 5.65 (on the 56-point frequency of use scale) for an individual with 20 or more years teaching experience, versus just 1.13 points on the frequency of use scale for a teacher who had been teaching one year or less. As a result, the predicted gains in frequency of use of standards-based instructional and assessment strategies are 8% greater for older, more experienced teachers than for new teachers. These findings are logical; in general, older, more experienced teachers have more to change than younger teachers who were perhaps trained to use teaching strategies in line with standards. As a result, it would be expected that individuals who have many years of teaching experience would demonstrate greater residual gains in their teaching strategies (following nearly a year of involvement in a change initiative) than their less experienced counterparts.

This model accounts for almost one-third (28%) of the variance in residual gains in use of standards-based instruction and assessment. While 72% of the variance is still unaccounted for, these results are superior to those for residual gains in level and intensity of concern, which only accounted for 13% and 18% of the variance, respectively. This suggests either that the variables selected for this study are better predictors of changes in use of standards-based teaching strategies than of changes in concern or, perhaps, that changes in concerns about standards-based instruction and assessment are simply more difficult to predict.

Analysis of Teacher Interviews

One hundred thirty-two interviews with 69 teachers at five of the schools involved in this study were content analyzed. Each teacher was interviewed between one and five

times over the course of the academic year. A specific breakdown of the number of times teachers were interviewed is found in Table 21. At each site visit, the principal was also interviewed.

During the interviews, the researcher used a general interview protocol as a guide; however, additional questions and probes were used as needed. A copy of the interview protocol is located in Appendix 4.

Teacher interviews yielded additional data on many of the variables targeted in teacher surveys and generated from teachers' professional development activity logs. Summaries of teachers' perceptions of and experiences with relevant variables are presented below.

Table 21

Number of Times Teachers Were Interviewed

Number of Interviews*	Number of Teachers
1	31
2	14
3	11
4	9
5	6

* The total number of interviews in this column sums to 158, instead of 132. This discrepancy is due to the fact that some teachers were interviewed in groups of 2 or 3.

Implementation of Standards-Based Activities in the Classroom

For the teachers interviewed, the major characteristic of their involvement in the staff development project was the opportunity (and in some cases the mandate) to implement standards-based math instruction and assessment in their classrooms. This corroborates information provided by teachers in their Professional Development Activity Logs, where teachers indicated that the primary staff development activity in which they engaged during the year was carrying out standards-based activities in the classroom. In fact, 75 of the 89 teachers who completed Professional Development Activity Logs indicated that one of their staff development activities was implementing standards-based curricula.

Information regarding the use of standards-based math activities in teachers' classrooms was generally gathered via questions and probes including, but not limited to, the following:

- What is your role in this project? What are you doing as part of the project?
- Compare your math teaching this year to last year. How is it different/the same?
Why?
- Are you sacrificing anything in order to do this?
- What problems are you encountering with these new practices?
- What new information have you learned about your students as a result of your changes in practice?
- What can your students do differently now?

Over the course of the year, most teachers reported that they were putting aside their old math textbooks and focusing almost exclusively on implementing standards-based math curriculum. In particular, they indicated that they were using rubrics more regularly, conducting problem-solving activities on a daily basis, asking children to draw their answers to “see what’s going on in their heads,” asking students to explain their answers, and sharing more information with parents about the criteria used to evaluate their children.

Teachers reported that they were using rubrics and talking to students more frequently about rubrics. On the whole, teachers also reported that they were observing students more and identifying their specific problems and strengths. One teacher indicated that, for the first time, she could name every student’s strengths and that it had been quite a learning experience teaching via students’ strengths, rather than their weaknesses. Teachers also reported that they had begun giving students more time to respond than in the past.

Teachers also indicated that students are working and discussing more in groups. One teacher related that her role in the classroom had changed and that it was fun to listen to her students. Before, she used to teach a lesson and then the class would do a worksheet on what was taught. Now, there was more investigating, observation, discovery, and teamwork in her classroom. The teacher stated that it was motivating for her to watch students come upon the answer she had hoped they would discover and that she preferred teaching this way.

The majority of teachers felt that their math teaching had truly changed. They indicated that they had slowed down to listen to children's responses and that they spent more time writing and discussing math than ever before. They also stated that they were extending activities to real life math problems that students needed to solve in school. Teachers also felt that "it was more fun to teach this way."

The utilization of new curriculum also inspired some teachers to collaborate more closely than in the past. For example, some teachers were working in pairs, having their students complete the same math activities and then meeting to observe and analyze student work. They found this exercise useful because "being able to talk about [student work] helps [them] reflect," particularly on the wide range of strategies used by various students to solve a single problem. Others had begun to team-teach and remarked that working with a colleague gave them the opportunity to walk around the classroom and observe their students more closely than usual (especially the "quiet" students). Collaborating teachers also stated that their collaboration activities compelled them to be more accountable. Each teacher knew that the other teacher was relying on her, forcing her to plan and implement what she had promised to do.

On the whole, teachers felt that the opportunity to implement standards-based math activities had helped them "see more" and delve more deeply into what students were doing. However, they recognized that they needed a management system to remember and record what they see children do. As a result, some teachers were working on locating and/or developing observation instruments/charts to record what they saw children do. Others had begun to look for assessment tools and observation sheets that

could be used to increase their consistency and objectivity in assessing different children.

This, in turn, had opened up another question that perplexed teachers:

“Here’s what you’re seeing. What are you going to do next and how is it going to work instructionally? You solve one question and open up 100 more. I’m glad I went there, but the more you know, the more you know you don’t know.”

Some teachers wrestled with the ideal balance between “skills” and standards-based math activities. For example, a kindergarten teacher reported that she still used the traditional math text because she felt that students needed experience with paper-and-pencil activities as well as standards-based activities. She stated that if students did not know how to write a “2” or a “7,” she could not ignore these skills or send those students to first grade. Another teacher reported that she still gave a lot of standard multiplication exercises. However, she indicated that this only took about 10 minutes per day and did not interfere with standards-based math activities.

Other teachers are more comfortable with the skills/standards balance. One teacher indicated that using standards-based math curriculum had actually helped her relax about this issue. She reported that her students could do number sentences without her directly teaching this concept because number sentences are embedded in standards-based curricula. She had seen students gain knowledge on their own, which she believed to be more meaningful to them than directly teaching skills. She stated that students were discovering concepts on their own, and her role was to put a name to them. For example, when students discovered that $10+0$ was the same as $0+10$, she explained that this was an example of the commutative property.

At the conclusion of the year, one principal remarked:

“At the beginning of the year, teachers probably would have liked someone to come in and tell them how to use [the standards-based math curriculum]. However, that person did not exist, so teachers tried it out on their own. This was probably the best way to do it because it is where and how the real learning takes place. Teachers are so used to listening to speakers, but this kind of professional development is entirely different—and valuable.”

While they were excited about the changes they were making in their own teaching, teachers were most enthusiastic about the effects they perceived that standards-based math instruction and assessment had on their students.

Teachers noted that the math terminology being used by students over the course of the year was more advanced than that of past students and that students’ writing in math (e.g., explaining their reasoning and communicating their solutions in writing) had improved. Teachers were also pleased with the math discussions they had begun hearing in their classrooms. Students were not only displaying a higher level of vocabulary, but they were more confident explaining math concepts and their math reasoning. As a case in point, teachers of lower elementary students were pleasantly surprised that their students understood and used terms such as “array,” “body benchmark,” and “multiples” in ways that they could not have done in the past.

Teachers were also impressed with the strategies students used and with the fact that students frequently employed more than one strategy to solve a problem. They also felt that students’ recall had increased, as a result of explaining and writing about math.

Some teachers believed that students actually understood mathematics more deeply using standards-based curriculum. For example, one teacher stated that her students understood multiplication better than any class she had ever taught. She was surprised about this, because she had not thought that using standards-based math

curriculum would be “enough” for this deep understanding. One teacher described the significant changes she had observed in her students, stating:

“I really pushed the kids hard, but they’re getting the hang of it. They have changed their whole perspective. Before, they were doing lots of skill work. Now they’re writing things out, explaining a problem, and writing it out in longhand. They have the concept of graphs, range, median, mode—they really have it. They’re improving their vocabulary, too. It’s been a long process to get them to this point. That period of time was really frustrating, but they’re thinking in terms of solving multi-step problems, and they’re starting to understanding that there’s something deeper.”

One teacher indicated that kindergarten students’ skills were much stronger than in the past. With the old series, kindergartners needed to know facts up to 12. This year’s students were “already way beyond that.” Additionally, her kindergarten students appeared to understand place value (which was not stressed in traditional kindergarten math) much better this year. This teacher also commented: “Kids reach beyond because they are free to do so. They are not held back by the book’s expectations.”

Teachers were also very enthusiastic about children’s increasingly positive attitudes toward math. They also remarked that students had come to believe that everyone could do math and that understanding concepts is more important than speed. One teacher reported that she had children who would be eligible for remedial math services, but these students are not ashamed or afraid to explain their reasoning in front of the class. One teacher noted:

“It’s a different atmosphere. Years ago, kids like that would never speak in math. They would be afraid they have wrong answer. Now, the rest of class will correct them nicely, and it becomes a learning experience. The class accepts everyone’s contribution even if it’s not right. They talk about why it’s not right. Also, there is a willingness to work with everyone. They used to only want to work with kids who got an A in math. They are more confident in their ability to do math than

they've ever been. They feel like mathematicians and when they feel that way they are mathematicians. It's all attitude.”

Teachers also remarked that students had become more comfortable working together in teams and discussing and defending their math reasoning. Additionally, students had learned that a single problem could have 2 or 3 right answers. An artifact of this discovery was that students had acquired a greater tolerance of the methods other people used.

Teachers also reported that their experience using new math activities had raised their expectations of students. One teacher stated:

“There are some things I think won't work--the kids won't be able to do it. Then, the kids can do it. It forces you to give them more challenging, mind-stretching activities. They're there. Maybe you think it isn't going to work, but it's in there [the new curriculum], so I'll try it. Then I just stand back and watch. In all the years I've been teaching, activities like that didn't work. Because of the hands-on, the communication, it's working now.”

Through their use of new math strategies and materials, teachers felt that they had learned that students were much better problem solvers than they had previously thought. They reported that students are able to perform better and do more when they are given a broader range of activities than when, for example, they are asked to “do the same thing 35 times.” One teacher explained, “Now that we've opened the door, they can tell us so much more than we gave them credit for.”

Interestingly, the small subset of teachers who resisted implementing new, standards-based mathematics curriculum displayed some aspects of low self-efficacy, particularly in their willingness to persist in the completion of a task. They also did not believe that their students were capable of succeeding at more challenging math tasks.

“[The new standards-based math curriculum] is uninteresting. The group projects don’t always work, and some investigations go on and on. I don’t want to spend so much time on one thing. Also, the topics are too advanced, and the kids aren’t ready. I have to bring it down to their level.”

“I can’t spend too much time on math because the kids get too frustrated. They have to know what they’re doing before you can let them go off and do their own thing.”

“In Investigations, skills are being neglected to teach the concepts. But kids need to know what the number ten means. Especially with ESL kids, they need more repetition and skill development than I’m doing this year [with the new curriculum].”

It should also be pointed out that the teachers alluded to above tended to work in a “vacuum,” isolated from lead teachers and the local consultant. They were ill-informed of the nature or goals of the professional development project and did not actively participate in any group staff development activities.

Data from interviews suggested that implementing standards-based classroom activities helped teachers learn new things about teaching and students’ ways of thinking. In addition, it appeared that teacher self-efficacy (the willingness to see a task through and the belief in one’s ability to perform tasks in the face of adversity) was linked to teachers’ willingness to put in more than a cursory effort at using new curricula. Teachers who displayed traits associated with low self-efficacy found their new standards-based math curricula uninteresting and had little desire to incorporate it into their teaching repertoire. These findings support the statistical results in this study suggesting that self-efficacy is a significant predictor of intensity of concern about standards-based mathematics. As stated previously, these conclusion suggest that it

might be advisable for planners of staff development projects to address self-efficacy issues with teachers.

Interview data also indicated that teachers learned valuable lessons about themselves and their students in the process of implementing standards-based activities in their classrooms. In addition, statistical results revealed that this activity interacted with action research to significantly predict residual changes in level of concern. However, experimentation with new activities in the classroom was not associated with gains in the frequency of use of standards-based strategies in the classroom, suggesting that this tactic may interact with action research to change attitudes but not practice (at least in the short term). Therefore, it is essential that staff developers employ other strategies and address other issues, as well, such as examination of student work, self-efficacy “training,” and organizational development focusing on improving principal supportiveness.

Work with Educational Consultants

Data on teachers’ views of their work with the educational consultants who had been assigned to support them were gathered via probes such as the following:

- What do you see as the role of the local consultant? What do you hope to “get” from him/her?
- How much contact have you had with your consultant? What have you done with him/her? Why haven’t you had contact with him/her?
- What have you learned with/from the consultant? What have you applied? How did it work?
- What more would you like to do with the local consultant?

- Approximately what percentage of teachers are working with the consultant?

Interviews focusing on the role of the local consultant generated interesting information regarding high expectations teachers had for them. While teachers recognized that a great deal of informal learning about teaching and about their students had occurred “on the job,” through the process of implementing standards-based instruction and assessment, they expected most of their “formal” learning to take place via the educational consultant assigned to their schools. In fact, teachers had extremely high expectations of their local consultants, who tended to be university professors or retired teachers or administrators. They expected them to perform a number of duties and possess a great deal of expertise, not only in mathematics, standards and standards-based instruction, but also in facilitating groups, modeling lessons, training teachers to use new curricula, guiding the action research process, and identifying/obtaining resources.

Teachers and principals recognized that educational consultants played additional roles, as well. Both groups of people saw the consultant as a bridge between the principal and the teacher. In addition, they stated that it was the consultant’s responsibility to help structure the staff development project in each school and provide direction for the project. For example, one teacher noted that an important role of the consultant was to “take some responsibilities off the plate of the principal.” One principal also stated that she had permitted the local consultant to assume some of “her” role in the project. Consequently, some teachers recognized that consultants took some stress off principals, “making [them] much, much more supportive.”

Teachers were unanimous, if not a bit unrealistic, about the primary role of the local consultant:

“If you have the right consultant at school, that’s probably one of the most important parts. She’s the glue that holds it all together. She’s helpful, always available, offers to help. She’s always there and never late.”

As the “glue that holds it all together,” educational consultants tended to be viewed as directly responsible for the success or failure of the staff development project at their schools. Furthermore, most teachers, including lead teachers, tended to “sit back and wait” for the consultants to “do something.” For example, when a consultant cancelled a meeting at one school, causing much displeasure among teachers, lead teachers refused to facilitate the meeting, stating “The meetings are HER meetings.” Another lead teacher explained her rationalization for waiting for the consultant to lead, rather than taking on a leadership role herself :

“There’s not much direction this year... It’s the responsibility of the consultant to do this, but the direction should come from the consultant. She went to Michigan and gave the teachers an activity to do, we did it, and we came back and discussed it. Teachers said, ‘Finally, we have something to do.’”

It was also extremely important for teachers to feel that their consultant considered working with them to be one of his/her top personal priorities. For example, many teachers were aware that most consultants had full-time jobs elsewhere and consulted for the project as something “extra.” Understandably, teachers did not want to feel that their consultant’s focus was somewhere else. They especially did not appreciate feeling that they had to “make a fuss” to get a meeting with their consultant. On the

contrary, one teacher stated, “Joan’s [the consultant’s] job is to be a consultant. This is a priority. This makes the whole model successful.”

Furthermore, teachers tended to feel that their relationship with their local consultants was at the same time professional and personal. When they did not see their consultant for a period of time, they worried that they had offended him/her in some way. When consultants had to leave the project for some reason or another, principals and teachers truly felt abandoned. When one unlucky school experienced the turnover of two local consultants, the principal explained, “Jill left us completely, and then Leslie quit on us, too.”

When teachers did not have frequent contact with their local consultants, they described the staff development project as “stagnant” and “null.” They added that the momentum of the project slowed considerably between meetings with their consultants. Over and over again, teachers reiterated that the consultant was one of the most important aspects of the project and that his/her presence and involvement was instrumental in keeping teachers focused on the goals of the project.

When teachers did spend time with their consultants, they became “energized.” They appreciated practical suggestions offered by consultants and frequently described consultant-led discussions as “good thinking sessions.” Overall, teachers felt that the most helpful activities carried out by their educational consultants were those instances when consultants modeled lessons in teachers’ classes and held subsequent debriefing session with them. Teachers indicated that it was useful to see things being implemented, rather than simply hear about how they should be done. They also felt that they learned

more about themselves and their students through these experiences. For example, one teacher said that the way the consultant probed the students for questions at the beginning of her lesson made her realize that she needed to direct students less and do more probing.

Interestingly, however, very few teachers had the opportunity to participate in this type of coaching experience with their local consultant. Of the 89 teachers whose Professional Development Activity Logs were analyzed, only 9 teachers had the chance to observe a consultant model a lesson over the course of the year. Furthermore, the average amount of time these 9 individuals observed consultants modeling classes was 30 minutes over the course of 7 months. Perhaps this explains why the consultant variable was not a significant predictor of residual changes in teachers' level of concern, intensity of concern, or frequency of use of standards-based strategies in the classroom.

Of course, it is possible that the teachers who participated in interviews and raved about the impact of consultant modeling were not the same people whose logs were analyzed. Nevertheless, findings from teachers' Professional Development Activity Logs do suggest that few teachers had the opportunity to watch someone else teach.

Curriculum Development

In the interview protocol, there were no probes focusing directly on the degree to which teachers developed or modified their own standards-based curricula. Rather, information regarding curriculum development surfaced during conversations focusing on teachers' efforts to implement standards-based activities in their classrooms. Teachers who undertook curriculum development tasks tended to be the same teachers who

encountered difficulties implementing new curricula or who felt that their curricula did not meet the needs of their students.

Approximately one third of the teachers in this study (31/89) elected to develop or revise their own standards-based math curricula. The majority of teachers, however, were immersed in the process of implementing new standards-based math curricula adopted by their schools and/or districts and had neither the time nor the inclination to write their own curriculum. However, teachers at two schools felt that the standards-based curriculum in place at their schools needed modifications and/or additions.

At one school, teachers, who were in their second year of the implementation of a new math curriculum, were unhappy with the quality of assessment tasks in their new curriculum. At the beginning of the school year, they thus set about the mission of developing their own standards-based assessment tasks. Very soon in the curriculum development process, however, teachers realized how difficult it was to develop high quality math tasks, admitting that they “had taken on a huge chunk” developing their own standards-based math tasks. One teacher at that school admitted, “My objective had been to develop a task, but that’s not my expertise.” As a result, they quickly decided to refocus their efforts away from developing tasks and toward the identification and field-testing of exemplary tasks that had been developed elsewhere.

At a school for deaf children, teachers reiterated the fact that the language level and much of the background knowledge/experience assumed by the authors of their new standards-based math series were inappropriate for deaf students. They thus embarked on the task of modifying the curriculum they were charged with implementing. Teachers

found the curriculum modification process extremely time consuming and frustrating, causing more than one teacher to ask:

“If you have to rewrite everything, do you want to use this program?”

By the end of the year, most of the teachers at the school for the deaf had concluded that standards-based math was completely inappropriate for their student population.

This may explain the fact that stepwise regression analyses revealed that involvement in curriculum development had a negative effect on residual gains in teachers’ intensity of concern about and frequency of use of standards-based instruction and assessment in math. Interview results suggest that the task of developing curriculum was so daunting that teachers effectively gave up on using standards-based math curricula. In turn, their enthusiasm for this reform waned substantially.

One feature of the curriculum development process at the two schools where teachers attempted to develop curriculum was the fact that, in both schools, teachers undertook the curriculum development task *alone*. Few had prior experience in curriculum development, most had only a limited knowledge of standards-based mathematics, and none received assistance with the curriculum development task. At the school where teachers began writing assessment tasks, the local consultant’s primary role was modeling lessons and training students and teachers to use new calculators. At the school for the deaf, no consultant consistently assisted teachers in any capacity. These factors no doubt clouded teachers’ curriculum development experience and suggest that curriculum development should not be undertaken without some sort of expert guidance or support.

Workshops

Fifty-six percent of study participants (50/89) took part in some type of workshop training activity over the course of the study period. Lead teachers attended a series of 3 weekend workshops; in addition, lead and participating teachers participated in several district and state-sponsored workshops. For the most part, these district and state workshops focused on standards, standards-based teaching, rubric development, and the nuts and bolts of implementing the math curricula that schools had adopted.

Data regarding teachers' attendance at and perceptions of the utility of workshops were gathered via interview probes such as:

- Have you attended any workshops? Which ones? What did you learn? What have you applied? How did it work?
- What have you shared from your workshops with other teachers at your school? How did this sharing occur?
- Is there any way the workshop could have been improved? What do you hope to learn next time?
- Aside from workshops, in what other ways has your school focused on standards in math teaching?

An examination of teacher interview data reveals that the old adage, "You can please some people some of the time, but you can't please everybody all of the time," applies to teachers' perceptions of the impact and usefulness of workshops as a staff development activity. Teachers attending the same workshops left these sessions with completely different reactions to them.

For example, some lead teachers attending project workshops viewed these sessions as practical and down to earth. In addition, they felt that their opportunity to “*drop everything and read*” professional development literature (i.e., DEAR time) during the workshop was “huge” in terms of the usefulness and benefits of professional reading.

On the other hand, many teachers found the workshops too theoretical and yearned for more “how to” information. Others still were disappointed at the use of workshop time for teachers’ own professional reading. Some resented using their weekend time to sit and “do nothing,” while others felt that they were being treated like children when they were given specific material to read for a set amount of time.

Similarly, some participants at a state-sponsored curriculum training workshop felt that there were too few opportunities for hands-on learning experiences. However, others stated that they would have preferred fewer hands-on activities and more of the theory behind what they were doing.

These results suggest that large professional development workshops were too broad to meet the individual needs of the teachers who attended them. In an effort to address a general level of teacher skill, knowledge, and learning styles, workshops failed to be relevant for many study participants. These findings also support statistical results in which attendance at workshops was not instrumental in predicting residual gains in teachers’ level of concern, intensity of concern, or frequency of use of standards-based strategies. In addition, these results corroborate the conclusions reached in previous research conduct by Sparks (1983, 1986) and Wade (1985), who asserted that training workshops alone had a minimal impact on teachers’ classroom practices.

Action Research

A significant component of the staff development under study consisted of teachers' initial forays into the area of action research. For the purposes of this study, action research was defined as the process of identifying "essential questions," or research topics, collecting data to address essential questions, analyzing data, and compiling final reports or projects to showcase one's action research results. Questions and probes regarding teachers' participation in action research were embedded in conversations regarding training workshops, work with educational consultants, and meetings with colleagues.

Slightly more than one third of the teachers in this study (36/89) engaged in any of these action research activities; furthermore, for those who did conduct action research, the mean amount of time devoted to this form of staff development over the course of the study was a scant 2.13 hours. In addition, lead teachers conducted almost twice as much action research as participating teachers (2.98 hours versus 1.63 hours).

A two-fold explanation for the low level of participation in action research exists. First of all, communication from lead teachers to participating teachers about the nature, goals, and process of the action research project was lacking. As a result, lead teachers tended to be the only ones who carried out action research. Second, teachers, who had no prior experience in action research, did not receive adequate support in *how* to carry it out. This, in turn, led to a minimal effort in the endeavor and a "just get it done" mentality.

Lead teachers, who were true novices in action research, learned what they knew about action research at three 2-day workshops which took place in October, February, and May. While these workshops did not focus exclusively on action research, lead teachers were nonetheless expected to take the little they had learned about the topic and transmit this information to the lead teachers at their schools. However, lead teachers found the whole subject nebulous and repeatedly complained that they had nothing to share with participating teachers at their schools following the lead teacher workshops. The result of this was that participating teachers knew very little about the action research aspect of the staff development project and, for the most part, only lead teachers conducted any form of action research.

Even at the end of the project, some participating teachers still said that they did not know if the school ever came up with an essential question. Others said that they knew an essential question was voted on at a school meeting but there had been no further discussion after the vote. Others were able to state the essential question but were not clear about what they were supposed to do. When queried about her involvement in action research, one participating teacher stated:

“I didn’t know I had to do anything more than record my experiences.”

Lead teachers were not completely at fault in their failure to include participating teachers in action research. Principals and teachers alike commented on the fact that they did not receive regular support in this endeavor. In addition, they were concerned that their local educational consultants tended to be “math people,” and not experts in action research. One principal noted:

“Teachers need more instruction in action research. They also need someone from the outside coming in regularly, say once a month, to help them through that. I’ve read enough of action research to know that someone usually does this.”

The frequency of support for action research seemed to play a key role. For example, lead teachers at one school explained that the infrequency of their meetings with their local consultant (approximately one time per quarter) caused their action research projects to “take a back seat” to many other school issues that appeared more pressing and urgent.

The single study school that appeared to meet some success with action research was rather unique. This school was extremely small (6 teachers), and all teachers were lead teachers. As a result, each person attended lead teacher workshops where action research was presented, and the problem of communicating with participating teachers was irrelevant. In addition, this school’s local consultant was very responsive and involved in what teachers were doing.

In the spring of the study year, these teachers organized and participated in a 1-1/2 day (late Friday afternoon through Saturday) retreat at a local campground. At this retreat, teachers looked at student work, discussed their essential question, and drafted their final project. While teachers felt that the retreat was very useful in terms of what was produced, they commented that this was partly due to the fact that they took time early in the retreat to make sure that everyone was approaching the retreat from the same mindset. They did not proceed to work on their essential question until they were “all on the same page.” While they were clear that the time they spent together was intense and productive, these teachers were also grateful to be away from the school building, feeling

that this allowed them to focus more clearly on the tasks at hand. Because this retreat took place outside school time, teachers commented that they felt less stress than they did during typical in-school staff development activities. In fact, one teacher said: [During the retreat], I was more relaxed than I had ever been at a workshop. “

These teachers also realized that working on their essential question yielded experience and knowledge that empowered them as teachers. The process of conducting action research also caused certain teachers to reflect more, particularly on their students’ thinking. They remarked:

“You have to experience it. People can tell you what to do because they’re experts and know, but until you experience it on your own, you don’t grow.”

“Going through the entire process was the most valuable professional development experience. It was a long process, but it will be lasting.”

“I got a lot out of the whole experience as much as I was tearing my hair out. I’m pleased with the final product and I learned a lot. I learned a lot more about the standards. I learned lots more about the thinking behind the math behind certain tasks. And I learned a lot about students’ thinking.”

While few teachers in the project actually participated in action research, it seemed to be a powerful experience for those who did. In addition, stepwise regression analyses revealed that participation in action research interacted with the implementation of standards-based classroom activities as a significant predictor of residual gains in teachers’ level of concern. In fact, the more involved teachers were with the implementation of new curricula, the higher the association between action research and residual gains in level of concern. As such, action research is a strategy that change agents and staff developers might want to employ more extensively with teachers, but not in isolation. The findings of this study suggested that action research alone was not

highly associated with residual in level of concern; rather, it appeared that teachers must be simultaneously making changes in their teaching. In addition, interview results indicated that intensive collaboration and follow-up with an expert consultant enhanced the effectiveness of action research.

Meetings/Discussions with Colleagues

Based on data collected during teacher interviews, it was possible to conclude that teachers felt that frequent meetings were essential to their professional growth, even if statistical results did not convey the same message. Over and over, teachers lamented the fact that they were not able to collaborate more with colleagues in meetings and/or group discussions. For many teachers, scheduling restrictions and a lack of communication about the project from lead teachers prevented teachers from meeting together.

Teachers at those schools that did set up regular opportunities for teachers to share (e.g., monthly project meetings) found these experiences to be useful for getting new ideas and finding out what other teachers were doing in their classrooms. They also felt that these meetings were contributing to everyone “being on the same wavelength” within the school building. They pointed out that having scheduled meetings kept them focused on the project and more aware of the teaching strategies they were using in their classrooms. In fact, the only consistent negative feedback regarding regular project meetings was that they were not held often enough!

According to the majority of teachers, the most valuable use of meeting time was viewing and analyzing videotaped lessons taught by their colleagues. During these meetings, teachers watched and discussed the videos, observing individual students,

noting how students interacted, and making note of the strategies used by the teachers in the videos. Lead and participating teachers alike indicated that the use of video allowed them to perceive details about students that they would not have noticed otherwise. Teachers reported that this activity also showed them the strengths and weaknesses of their peers and reassured them that all teachers were “in the same boat” with regard to their efforts to change their teaching and improve students’ skills. Teachers also stated that they gained valuable ideas from each other and had the opportunity to view teaching and student work in a different light. In fact, many teachers who were initially uncomfortable with the idea of being videotaped and having others comment on their teaching became increasingly comfortable with the activity and began to express interest in being taped.

In conclusion, teachers’ interviews made it evident that teachers considered meetings a valuable use of staff development time when they were well-planned and centered on an activity that was closely linked to teachers’ work in classrooms. Teachers were clear that “gripe sessions,” while they might be effective in releasing stress, were not useful in terms of their professional development. In addition, teachers averaged 13 hours of meetings over the study period, making meetings the second most popular staff development activity utilized by study participants. (Only standards-based classroom activities were implemented more frequently than meetings. Despite the high value teachers placed on opportunities to meet and the time that teachers actually spent in meetings, this staff development activity did not appear powerful enough to predict changes in teachers concerns or teaching practices. Either meetings are over-rated, or the

quality of the meetings which teachers attended was not sufficient to affect change in teachers.

Examining Student Work

During their involvement in the staff development project, teachers began looking at student work in ways that they had not done before. This, in turn, encouraged many to reflect more closely on their own teaching and the rationale behind the way they did certain things in the classroom. Teachers employed different techniques for examining student work. Some saved student work sampled and compared student performance over time. Others deliberately observed students in the process of completing math tasks and took note of relevant details surrounding their performance. Other teachers videotaped their students in class and viewed them later, closely scrutinizing the strategies individual students employed. Many also scored student work using rubrics that they or others had developed.

This reflection on student work often showed teachers things that they did not realize about their students. An example of the learning that took place included:

“My assignment was to view kids and listen to their accountable talk. I saw that the terminology kids are using is more advanced than I had ever thought. I was also impressed with the strategies they use and that they complete projects with more than one strategy. I never had seen that before. Also, they branch off to other forms of math, topics within a single topic.”

Teachers also reported that they had learned new ways of solving problems from their students, causing them to realize that multiple approaches exist for problem solving. Consequently, they not only gained insight into how their own students thought but were also moved to reflect on their own ways of working. Many teachers subsequently felt

that their teaching had improved, as they had begun to examine what “students [did] not understand and use what they [did] know to focus on where they needed to go.” For example, in the past, lead teachers reported that they had focused on students’ getting the right answer. Now, they were focusing more on children’s math reasoning and examining students’ thought processes.

Perhaps more than anything, the experience of analyzing student work raised teachers’ expectations of their students. By looking at how students approached tasks, many teachers stated that they realized that students were much better problem solvers than they previously thought. They also learned more about the learning styles of individual students and ways to help diverse students succeed.

One principal summarized the benefit of close examination of student work on her staff:

“Teachers got to reflect on student work and on their own ways of working. They are asking themselves, “Why am I doing this? Is this worksheet worth my time?” In the past, they set goals but didn’t evaluate or reflect on them enough. The project taught this very experienced staff the value of doing this.”

A teacher reiterated this view:

“The project forced me to look at a new way of teaching and do some “kid watching.” I underwent a metamorphosis.”

Based on teacher interviews and statistical analyses, it appeared that the effects of examining student work in new ways was powerful in terms of the influence it may have on teachers concerns about standards-based instruction and assessment. While teachers in this study averaged just 1.23 hours on the student work variable, it was nonetheless a significant predictor of changes in teachers’ level of concern. In addition, those teachers

who engaged in this type of staff development were very conscious of the fact that the opportunity to study students more closely had altered their notions of teaching and learning.

Communication and Organizational Climate

As shown through teachers' experiences with consultants and action research, the presence or lack of direct face-to-face communication among the parties involved in this staff development project appeared to influence the success or failure of various development ventures. For example, a lack of communication among lead and participating teachers limited the access of all teachers to potential professional growth via action research. In addition, limited communication among local consultants and all teachers probably influenced the impact of action research efforts on teachers' every day practices.

In this study, some teachers felt informed about the project, while others did not. On the whole, however, most teachers (lead and participating) wished that they had more opportunities to meet together and discuss the project, both within their own schools and across schools. In addition, all teachers recognized the importance of regular communication. One teacher summarized the importance of communicating with colleagues:

“When you talk all together, it gives you the opportunity to see the value of what you're doing at your level and how it applies to the next level. What happened before me is important, I am important, and I extend what is already introduced.”

On the whole, lead teachers met regularly with each other. In few instances, however, did lead and participating teachers come together as a large group to meet about

the staff development project. This situation did not occur because of a lack of interest on the part of participating teachers; on the contrary, participating teachers wished that they had more frequent, planned opportunities to meet with each other and with lead teachers. When meetings did occur, they were brief and spur-of-the-moment. One participating teacher said that even the chance to meet for a half an hour every 4 weeks would be helpful and would generate more widespread interest in the project.

Exceptions to this general lack of communication were two teachers who were regularly having their classes work together on math. These teachers were clearly seeing the benefits of their collaboration. One teacher said:

“When you work with another teacher, you have to force yourself to be prepared. Knowing that I meet with Cathy every Wednesday and doing an activity every Wednesday afternoon forces me to take the time to do it. I’m doing a lesson or two faithfully every week. I’m surprised to see how much we’ve done.”

In addition, a lack of communication was linked with organizational climate, in that schools with a low degree of communication also contained isolated sub-systems of teachers. Indeed, the correlation between communication and organizational climate in this study sample was .67 ($p < .05$). Of course, it is not possible to conclusively determine which came first: poor school-wide communication or a low quality organizational climate. For example, did isolated sub-systems of teachers exist in some schools because school staff could not or would not communicate? Or, did poor organizational climates (i.e., isolated sub-systems of teachers, weak channels and procedures for dissemination) lead to low degrees of school-wide communication?

Throughout the staff development project, participating teachers who rarely had the opportunity to meet with lead teachers, local consultants, or others involved in the

initiative indicated that they felt “way out in left field” and ill prepared for what is being asked of them in terms of the project. For example, one 5th grade participating teacher explained that she felt very isolated in the project (she did not even know who her lead teacher was) and stated that she wished teachers at her school could sit down as a team to “throw ideas off each other.” She wished that there were grade level meetings and that some team teaching could be done.

In fact, there was a subset of participating teachers at each school (except the small school comprised of just 6 teachers) who were very disconnected from the project. They did not know who the lead teachers at their school were, they had never heard of the essential question/action research aspect of the project, they had little to no contact with the local consultant, and they did not know what was required of them through their involvement in the project. On the whole, they equated the staff development project with “doing [the new math curriculum].” Furthermore, a theme repeated by these teachers was that they felt that they were not learning anything this year. Finally, they did not see a solution to this situation, as many reported that the organization of their schools was “not set up to share and talk.”

During the final set of interviews for the project, even lead teachers concluded that there had not been enough communication this year between lead and participating teachers. Citing time constraints and scheduling conflicts, lead teachers admitted that they were the only ones this year “who really knew what was going on.” They added that participating teachers had been better informed last year due to their attendance at mandatory participating teacher meetings. With last year’s structure, participating

teachers had more knowledge of what was expected of them, even if they were not able to meet with lead teachers.

One participating teacher summarized the sentiments of many of the participating teachers who were interviewed:

“Working Wonders started off good. After February, what happened? No one met, there were no schedules. We have no idea what lead teachers do or what they're supposed to do for us. There were no notices about meetings, but if you're a teacher who's willing to make changes, you'll use [the project] as a background, you'll be more into using math for problem solving, etc. I found I could do certain things [even without any support]. I tried to think of things to use from [the project] in my skills work, like problem solving and the discourse of math. I tell students they want to sound like mathematicians. [The project] made me more focused on these little things that make kids not afraid of math. I did what I needed to do.”

In addition, to indicating the degree to which isolated sub-systems of teachers are present in a school, the quality of a school's organizational climate also includes the strength the school's channels and processes for the dissemination of new ideas and information, as well as the organization's openness to innovation. As such, other factors that indicate a high quality organizational climate can include class schedules that accommodate new teaching and learning practices and the availability of teacher and student resources associated with a new innovation. Likewise, a poor organizational climate may be marked by an absence of these factors.

Teacher interviews revealed that teachers working in poor organizational climates were caught up in the day-to-day struggle of implementing new standards-based math curriculum and encountered too many impediments to make substantive changes in their teaching.

For example, teachers in some schools reported that they did not have all of the resources that they need to use their new math curricula efficiently. More than one teacher did not have a copy of the text book that she was supposed to be using. Other teachers expressed dismay at the amount of preparation that each math task required and the fact that so much photocopying had to be done to compensate for the fact that there were no student books available. Typical teacher sentiments were expressed below:

“I have to teach myself the approach first and make sure I have the right amount of manipulatives. If I don’t have them, I don’t do it [the math investigation] because I’d have to buy them.”

“I don’t have a copy of the New Standards. There’s just one copy for the 2 buildings... We’re in a very unique situation. We’re trying to do [the new standards-based math curriculum] but we have no books or manipulatives. There’s just one box of materials that isn’t even a full classroom set, but we have to share it among 5 classrooms. We’re supposed to get more next year. I know it will be a very difficult year.”

In addition, teachers stated that many new math activities required the use of overhead projectors, of which few were available. One teacher stated, “I keep seeing obstacles, and I don’t know how to get over them.” This lack of appropriate resources certainly hindered the dissemination of the new math curriculum.

Furthermore, some teachers recognized that the structure of the typical school day did not allow time for a different kind of math instruction. In fact, certain teachers indicated that they did not even have one hour per day for math. With a scheduled daily math period of less than a full hour, one teacher said, “The program requires a lot of ‘hands on.’ You sometimes just get set up and have to break up.” One teacher lamented that the school swimming program cut into math time, reducing math time to 1-1/2 hours

per week. In addition, this 1-1/2 hours was broken up, so that students only had 15 minutes of math on certain days. The lack of time to experiment with new math teaching practices led to increasing skepticism (on the part of some teachers) about the potential success of the entire staff development project and, logically, little time to explore with different ways of teaching math.

Where isolated sub-systems of teachers existed, many teachers did not participate in staff development activities that might have changed how they approached the teaching of math. In most cases, these teachers would have participated in more intensive staff development if the opportunity had been extended to them; however, they were uninformed of potential staff development opportunities. In addition, class schedules that did not accommodate new teaching and learning practices were an obstacle that most teachers were powerless to overcome. Furthermore, this showed teachers that their schools' openness to innovation was little more than "lip service." Finally, a lack of teacher and student resources materials for implementing standards-based instruction and assessment further reduced the likelihood that teachers would actually implement new math teaching strategies, as procedures for disseminating this innovation were extremely weak.

With regard to the above features (i.e., rigid class schedules, a lack of resources, and poor communication about the initiative), the degree of organizational climate becomes somewhat indistinguishable from the degree of principal supportiveness, variables that interacted together to significant predict changes in residual gain in level of concern. At the school level, principals are held responsible for disseminating

information, reorganizing existing structures to support change, and obtaining the resources needed to make change possible. As such, it is clear why the organizational climate and principal supportiveness variables were found to be so highly correlated ($r=.59$, $p<.05$). These two concepts are closely linked and play a potential role in forecasting changes in teaching practice.

Quality of School/Community Relations

With regard to the implementation of an innovation, the quality of school/community relations is partly determined by the degree to which parents and the community at large are aware of and involved in the innovation. Previous research (Manning, 1976) has suggested that successful educational innovations are more likely to be found in communities characterized by widespread parent and community involvement and support.

Interviews with teachers revealed that, for the most part, standards-based instruction and assessment did not enjoy widespread community support in the districts in which it was being implemented. In general, teachers indicated that parents were not particularly aware of or involved in standards-based reform in mathematics. One teacher stated:

“Overall, they [parents] are not aware. We have mentioned it [standards-based reform in math], and rubrics have gone home, but they need to be more informed. More parent-friendly literature needs to go out to them, so they can reinforce it at home. They need to know their role.”

However, some teachers indicated that parents were beginning to notice that the work their children are bringing home was different from what they themselves did as

children or what their older children did in school. Support from parents, teachers added, became greater as they increased in their understanding of the reasons behind the reform.

“A certain percentage are without a clue. Some parents don’t understand, so they’re frustrated with why children are getting that kind of work. But once you tell parents that it’s all over Rhode Island, they embrace it and realize that it’s more like real life than teaching math out of a book.”

“Some parents have asked about the homework. It looks different from what they’re used to. With [the new curriculum], parents *have* to get more involved with homework. The series is very parent involved. Children have to measure things around the house, ask their family questions. The parents like the “homelinks” [the family-focused homework]. They’re getting the idea that it’s different.”

Many teachers were concerned about the fact that middle and high school students “are still doing pages and pages of math problems” and that teachers at the upper grades had not bought into changing the way they teach math. These teachers were also sensitive to perceived parental concern about computation and the fact that parents saw that computation was emphasized in other schools (particularly middle and high schools) in the various districts. As a result, these teachers felt that “there has to be something uniform throughout” the system. In response to community concern, therefore, some teachers reported that they did “both.” they used standards-based math curriculum and computation-based materials.

There was general concern among teachers and the principal about what would happen to students who had learned math via hands-on, standards-based curriculum once they arrived at middle and high school. While they recognized that there was limited standards-based instruction and assessment in middle schools, these individuals felt that math instruction/assessment at the high school level was very traditional. As a result,

they remarked that it might have been better to start a project such as this one in high school. One teacher said, “If it came from the top down, it would be better for the students.”

In summary, teachers felt a good deal of pressure from their districts to implement standards-based math curriculum and raise student test scores. However, this did not translate to support from parents or the school district as a whole. Parents were relatively unaware of standards-based reform, although most lent some support once they were introduced to it. Teachers also tended to be responsive to parental concerns about math skills and computation; as a result, they added these features in to their math classes when they felt scrutiny from parents.

Teachers tended to view their districts’ focus on standards as disjointed and illogical. They felt immense pressure to implement standards but were acutely aware that teachers at the middle and high school levels were not being asked to change how they operated in the classroom. As a result, teachers were very worried about the fate of their students when they entered middle and high school.

In a sense, this aspect of school/community relations can also be viewed as a part of the school district’s organizational climate, if the concept of “organization” is extended beyond the individual school level and to the school system as a whole. While standards-based math was being pushed at the elementary level, the rest of the school system in most districts was not as open to innovation, severely weakening the dissemination of this change initiative. In this respect, it is plausible to conjecture that increased community support of standards-based reform in math might impact on the

frequency with which teachers employed standards-based math strategies in their classrooms (just as quality of organizational climate was associated with gains in frequency of use). With teachers aware that standards-based math did not extend beyond the elementary level, they felt pressure to teach traditional math in combination with standards-based math. Statements made during teacher interviews supported this fact. It is very logical, therefore, to assume that greater acceptance of standards-based math system-wide would free teachers to use standards-based teaching strategies more frequently.

Principal Directiveness and Supportiveness

Principal directiveness refers to the degree to which principals exercise control of schedules, meeting agendas, and school rules. The concept of principal supportiveness, on the other hand, conveys the degree to which principals share their knowledge, work and communicate with teachers, and offer constructive criticism. In addition, statistical analyses and teacher interviews revealed that principal supportiveness is inextricably linked to the quality of an institution's organizational climate. Unfortunately, teacher interviews did not reveal detailed information about the degree of principal directiveness and supportiveness at their schools.

On the whole, teachers were very reluctant to speak ill of their principals or suggest way in which principals could improve the implementation of the staff development project. This reluctance could signify loyalty on the part of teachers to effective principals or teacher fear that their comments would later be relayed to their

principals. In any case, teachers depicted their principals as non-directive and quite supportive.

In their interviews, teachers did not depict their principals as being particularly directive. This corresponds to survey findings, which indicated that the overall degree of teacher-perceived principal directiveness was lower than other school level scores. In fact, teachers felt that the push to implement standards-based math in their classrooms originated outside the school level, i.e., from the district, state, or beyond. As such, they did not view principals as directors of the staff development project or as the forces pushing them to change their teaching. Rather, teachers reported that they viewed their principals more as partners in the task of implementing new teaching practices and raising students' skill levels. Even in cases where teachers indicated that their class schedules did not support the implementation of new math practices, teachers did not blame their principals for this structure. Rather, they viewed this situation as something that their district would have to rectify.

Teachers also reported that their principals were supportive of their efforts to implement standards-based mathematics and the challenges this involved. Furthermore, principals who had recently changed roles from teacher to principal were viewed as particularly supportive. Teachers felt that these new principals understood where they were coming from, since they had just left the classroom themselves.

The sole occasion in which teachers let their resentment toward principals be known occurred in those instances when teachers did not have the materials or resources that they needed to implement standards-based mathematics. In these cases, teachers felt

that their principals were less than supportive. They did not appreciate being promised materials that never appeared or being expected to teach in a manner for which they were not prepared.

As noted earlier, principal supportiveness was a significant predictor of residual gains in the frequency with which teachers utilized standards-based teaching practices in their mathematics classes. These findings support previous research (Manning, 1976; Berman & McLaughlin, 1978; Gall, Fielding, Schalock, Charters, and Wilczynski, 1984) indicating that the support of the building principal for an innovation was directly related to the likelihood that staff would use project methods and materials.

Years Teaching

The findings of this study contradicted the results of the Rand Study (Berman & McLaughlin, 1977), which concluded that a consistent negative relationship existed between teachers' years of experience and teacher change. In contrast, this study found that more years of teaching experience were significantly associated with residual gains in the frequency with which teachers employed standards-based math instructional techniques. In reality, this was probably because individuals who had been teaching for the longest time had the most to change. After using traditional math methods for many years, most experienced teachers had to change many facets of their teaching repertoire to align with standards-based approaches. On the contrary, younger, less experienced teachers were less embedded in a single way of teaching; in addition, many were already trained to teach in a manner that corresponded with standards-based mathematics. As a result, they had less to change.

In general, teachers with many years experience were the ones to comment on how much the staff development project had changed their thinking and teaching. In addition, they commented repeatedly that what they had learned had renewed and invigorated them. In fact, more than one experienced teacher stated that she wished the project had come earlier in her career.

One 30-year teaching veteran remarked:

“My thinking has changed for the better. We’re all lifelong learners. This has been a new awakening. When you’ve been teaching for 30 years, you think you’re winding down, and then something like this comes up!”

A principal described the changes she had observed in her staff of very experienced teachers:

“Every once in a while, teachers need to be uplifted, excited about teaching. They have a renewed vigor for teaching math. They’re wondering how to fit it all in one day. They see the need for doing so much math, they’re probably teaching twice as much math as ever before. If that’s the only outcome [of the project], that’s good enough.”

Younger teachers tended to be upbeat about the project but did not share this renewed zest for teaching. Many reported that they already taught math in a manner consistent with standards and were involved in the project simply to look for more “tools” and ideas. Others felt that they had truly learned more about teaching and about students, but they nonetheless did not exhibit the excitement that the older teachers demonstrated. Rather, they were grateful for the opportunity to earn money and (in some cases) college credit for “something that was going to be mandated.”

What was unique, perhaps, about the findings of this study was the fact that older, veteran teachers were enthusiastic about change and *willing* to change how they taught

math. This does not match the stereotypical view many people hold of teachers. Experienced teachers took risks, tried new teaching techniques, and became more excited about the teaching of math than they had been in many years. Given the results of previous research on older teachers and change, these findings were surprising and may help dispel the notion that veteran teachers are not capable of adapting to change or trying out new things.

Summary of Results

In summary, regression analyses revealed that several of the variables under study were instrumental in predicting residual gain in level of concern, intensity of concern, and frequency of use of standards-based instruction and assessment. Teacher interviews supported these findings and offered concrete examples of how the variables affected teachers' involvement and experience in a change initiative focusing on standards-based reform in mathematics. Given that successful implementation of an educational innovation involves changing teachers' attitudes (i.e., their concerns) and their teaching practices, staff developers designing future change initiatives might want to consider planning staff development around the types of predictor variables that were associated with changes in teachers' concerns and practices:

- Personal variables: self-efficacy and number of years teaching
- Staff development activity variables: implementing standards-based classroom activities, action research, examining student work (in small teacher groups), and curriculum development

- Organizational variables: principal supportiveness (which has a great deal in common with organizational climate)

While the variables listed above cannot be said to have caused residual gains in teachers' concerns and practices, this study found them to be significant predictors of teacher change; as such, they are significantly associated with some of the changes that need to occur for a change initiative to be implemented.

From these results, one may surmise that an effective staff development effort around new classroom practices might be one that focuses on older teachers and includes activities centered on improving teacher self-efficacy, principal supportiveness/school (or district) organizational climate, as well as opportunities for teachers to examine student work and implement new teaching strategies (while simultaneously conducting action research). Based on these results, curriculum development is not recommended as an effective staff development strategy because it was found to overwhelm teachers and dampen their enthusiasm for change. Based on the results of this study, it is conjectured that combinations of these factors may be associated with residual gains in level of concern, intensity of concern, and frequency of use of new teaching techniques, all of which are necessary for successful implementation of an educational innovation.

In summary, Chapter 4 provided a description of the staff development activities, organizational features, and personal characteristics of 89 teachers who took part in a change initiative. The data also related these features to three measures of teacher change: residual gain in level of concern about standards-based instruction and assessment, residual gain in level of concern about standards-based instruction and

assessment, and residual gain in frequency of use of standards-based instruction and assessment strategies. Study results indicated certain staff development strategies, organizational features, and personal characteristics were significant predictors of the three measures of teacher change. A summary of these findings and a discussion of their implications is presented in Chapter 5.

CHAPTER 5

Conclusion and Discussion

In the first chapter of this dissertation study, the need for research regarding the degree to which staff development, contextual factors, and teacher characteristics might predict change in teachers' concerns about and use of standards-based instruction and assessment were outlined. In the second chapter, previous research on staff development and change theories were discussed. In chapters 3 and 4, the methodology underlying the present study was described, and results were presented which showed that staff development, contextual, and personal variables significantly predicted residual gains in teachers' level of concern, intensity of concern, and frequency of use of standards-based instruction and assessment. In this chapter, the procedures and results are summarized briefly, the contributions and implications of the study findings are discussed, the limitations of the study are presented, and finally, issues requiring further research are outlined.

Summary

Residual Gain in Teachers' Level of Concern

Stepwise regression analyses revealed that residual gains in level of concern were significantly predicted by teacher involvement in the following staff development activities: examining student work (i.e., saving student work samples and comparing student performance over time; observing and taking note of relevant details surrounding their performance; scoring standards-based assessments using rubrics; and setting benchmark papers) and implementing standards-based classroom activities while carrying

out action research. In fact, almost 20% of the variance in residual gains in concern was explained by a combination of the examination of student work variable and the standards-based activities/action research interaction variable. Teacher interviews supported these findings.

In addition to yielding information regarding the variables with the strongest modifying influence on residual gain in level of concern, the results of this study generated data regarding the number of hours of staff development that are associated with changes in level of concern. For example, statistical analyses demonstrated that, for every one-hour increase in activities involving the examination of student work, residual gains in level of concern increased by .11 (holding the effect of the classroom activities/action research variable constant). These results suggested that just 9.7 hours spent examining student work were associated with movement from one level of concern to a higher, more sophisticated one (e.g. from level 1 to level 2) on the Stages of Concern continuum in the Concerns-Based Adoption Model of Change.

Similarly, at “average” levels of implementation of standards-based activities (e.g. 17.22 hours over a school year), each additional hour of action research was associated with a .01 increase in residual gains in level of concern. Correspondingly, at “high” levels of implementation of standards-based activities in the classroom (e.g., 48 hours over the course of a school year), each additional hour of action research was associated with an increase in residual gains in level of concern of .40. These findings suggested that action research might play a beneficial role (in terms of residual gain in level of concern) when teachers are moderately to heavily involved in the implementation of

standards-based classroom activities. On the other hand, teachers who were not implementing these activities in the classroom yet undertake an action research project were not likely to experience residual gains in level of concern.

From the point of view of staff developers or administrators, staff development strategies in which groups of teachers look closely at student work and conduct action research are far more “cost effective” than simply handing teachers a new curriculum and expecting them to try it out. For example, the results suggested that just 9.7 hours examining student work or 11.76 hours of action research was associated with movement from one level of concern to a higher one (e.g. from level 1 to level 2). In contrast, the findings suggested that approximately 83.3 hours of time spent implementing standards-based classroom activities were associated with the same change in level of concern.

Another interesting aspect of these findings was the fact that no contextual or personal variables significantly predicted residual changes in level of concern. Information on organizational climate, school-wide communication, principal directiveness and supportiveness, and school size (contextual variables), as well as years teaching and teacher self-efficacy (personal variables) did not contribute to prediction of residual changes in levels of concern. This, perhaps, suggests that staff development initiatives focusing solely on helping teachers move from one level to another on a continuum of attitudes about change need only structure staff development activities to achieve this goal. If the goal is simply to have teachers think differently about an innovation, it is conceivable that a heavy focus on the school’s contextual features and on

personal teacher characteristics is not necessary. However, the results of this study indicate that contextual features and teacher characteristics are key in predicting changes in teachers' instructional and assessment strategies.

Residual Gain in Teachers' Intensity of Concern

The results of the stepwise analysis in which residual change in intensity of concern was regressed on the sixteen staff development, contextual, and personal variables in this study were slightly less clear cut. Results of this stepwise regression analysis revealed that teacher self-efficacy and involvement in curriculum development significantly predicted residual gains in intensity of concern. A closer examination of the results, however, suggested that the R^2 obtained in this analysis (.13) was significantly different from zero only at an alpha level of .10, rather than .05. This revealed that the possibility of a Type I error was slightly greater than would normally be desired. Because the purpose of this study was to *explore* the variables that might predict changes in teachers' concerns and practices, rather than *explain* teacher concern and practice, an interpretation of the model produced via this analysis was pursued. However, it was made clear that the likelihood of a Type I error in this analysis had increased to 10%.

With this caveat in mind, it was pointed out that increases in teacher self-efficacy were found to be associated with residual gains in intensity of concern. In addition, less involvement in curriculum development predicted residual gains in level of concern. Teacher interviews supported these conclusions, in that teachers who showed persistence in attempting to implement reform in their math classrooms (i.e., those with high self-

efficacy) were more positively intense in their concern about standards-based instruction and assessment. Furthermore, those who did not develop or modify standards-based math curriculum were, on the whole, more positive about standards-based reform in math.

As in the previous analysis, contextual variables did not significantly predict residual gains in intensity of concern. These findings, in combination with the previous results of the regression analysis on residual gains in level of concern, suggest that overall changes in concern are associated with participation in select staff development activities that emphasize implementing new activities, action research, examining student work and increases in teacher self-efficacy, regardless of school organizational climate (i.e., teacher cohesiveness, strength of channels of dissemination), the degree of principal supportiveness and directiveness, the quality of school-wide communication, or years of teaching experience.

Residual Gain in Teachers' Frequency of Use of Standards-Based Instruction and Assessment

Stepwise regression analyses revealed that almost 28% of the variance in residual gains in use of standards-based instruction and assessment was predicted by number of years teaching, degree of principal supportiveness, quality of organizational climate (a suppressor variable that enhanced the predictive ability of the principal supportiveness variable), and involvement in curriculum development. Again, teacher interview data supported these conclusions.

Clearly, prediction of residual gains in frequency of use of standards-based instruction and assessment proved to be more complex than it was for prediction of residual gains in level and intensity of concern. The results of this analysis revealed that a combination of personal (years teaching), organizational (principal supportiveness and organizational climate), and staff development (curriculum development) factors were influential in predicting residual gains in use of standards-based instruction and assessment. Given the fact that the sole staff development variable in the model, curriculum development, was associated with *negative* residual gains in use, its practical contribution to the prediction of residual gain in frequency of use was limited. In effect, therefore, staff development hardly played a part in predicting change in teacher practice. Rather, contextual and personal factors were key. These conclusions confirmed previous research findings by Manning (1976), Berman and McLaughlin (1977, 1978), Gall et al. (1989), Elmore (1996), and others, who found that factors of the school setting influenced the implementation of change more than any other variables.

As stated in Chapter 4, the findings of this study suggest that an effective staff development effort might be one that focuses on older teachers and includes activities centered on improving teacher self-efficacy, principal supportiveness, and school (or district) organizational climate, as well as opportunities for teachers to implement new teaching strategies, examine student work, and engage in action research. (Solitary, unsupported curriculum development efforts are not recommended.) Based on the results of this study, it is conjectured that combinations of these factors may be associated with residual gains in level of concern, intensity of concern, and/or frequency of use of new

teaching techniques, all of which are necessary for successful implementation of an educational innovation.

Contributions of the Research

Above all, the present study began the process of filling in some of the gaps in the staff development research. Prior to this study, few quantitative studies had been carried out on staff development models other than the training paradigm; in addition, the amount and types of staff development activities necessary to fulfill a given purpose (e.g., change teacher attitudes or teaching practice) were virtually unknown. Previous research on staff development has also tended to judge program effectiveness primarily by indices of teachers' satisfaction with a program, indicators of change in teachers' professional knowledge, or participation rates (Little et al., 1987; Guskey, 1995). Only rarely have changes in teachers' concerns about or use of an innovation been explored. In contrast, this study focused on residual gains in concern about and use of standards-based instruction and assessment, outcomes that are likely to impact classroom practice and student achievement.

This study explored the nature of several staff development strategies, teachers' reactions to and experiences with various forms of staff development, and the utility of predicting changes in teacher attitudes and practices based on the amount and type of staff development participated in. Preliminary findings suggested that certain types of staff development predicted teacher change more than others. The implication of this finding, of course, is that some staff development strategies are more effective than others at changing teachers' concerns and practices. In addition, study findings identified

the number of hours of staff development that were associated with given amounts of change. Both of the above points are topics that staff development researchers have long attempted to quantify.

This study was also the first of its kind to use statistical techniques to predict teacher change from a combination of staff development, contextual, and personal variables. In much of the previous research on staff development and teacher change, statistical analyses were limited to tests of group differences or simple correlations. This study was further strengthened by the fact that it incorporated both quantitative and qualitative research techniques, in contrast to most studies of teacher change, which tend to focus on one approach to the exclusion of the other. In most instances, findings yielded via quantitative methods validated qualitative findings, and vice versa, lending considerable credibility to the overall conclusions.

In summary, this study has contributed to the research on staff development and teacher change in a number of ways. First, it has offered new models for predicting changes in teacher concerns and practices. Second, this study has attempted to quantify the amount of staff development that is associated with given changes. Finally, the study has validated previous research suggesting that changing the way people teach requires more than staff development; rather, organizational, administrative, and personal factors (organizational climate, principal supportiveness, and teaching experience) may be key to effecting and sustaining changes in instructional practices.

Implications of Study Findings

Based on the findings that key staff development, contextual, and personal variables were associated with desired teacher outcomes (i.e., changes in attitudes and reported teaching practice), various recommendations are suggested for change agents and administrators who are planning or implementing change initiatives. In particular, these recommendations center on the following points: providing opportunities for teachers to examine student work, conduct action research, implement new types of activities in the classroom; not pressuring teachers to develop/revise curricula; increasing teacher self-efficacy, and improving principal supportiveness.

Statistical analyses indicated that increases in teachers' level of concern about standards-based instruction and assessment (based on the CBAM model of change) were associated with just over 9 hours of time spent examining student work. In addition, teacher interviews revealed that taking the time to observe student work closely was a powerful experience; they learned new things about how their students , increased their expectations of their students, and reflected deeply on their own teaching practice. Based on the findings of this study, therefore, it is recommended that staff developers and administrators incorporate this valuable staff development activity into their change strategies. Specifically, staff developers should encourage and guide teachers in the process of examining student work, while administrators should ensure that time and structure is built into teachers' schedules to carry this out.

Opportunities to examine student work that teachers in this study found valuable included saving student work samples and monitoring student performance over time,

videotaping students at work and viewing them later to identify their learning strategies, scoring student work using rubrics and discussing their scores with colleagues, and observing students at work and keeping notes on relevant aspects of their performance. Based on feedback from the teachers in this study, it is recommended that these activities be included in future change initiatives, especially those focusing on standards-based reform in mathematics.

Additional staff development strategies are recommended for helping teachers examine student work more closely. For example, it is recommended that teachers at the same grade level (or across grade levels) collect pieces of student work, score the work according to a rubric or standard, and 1) meet together to discuss their scoring, 2) document why the work did or did not meet certain criteria or standards, and 3) arrive at a consensus as to what constitutes work at various levels.

A second recommendation is that staff developers ensure that action research is carried out in conjunction with changes in teachers' everyday instructional practices. As presented in Chapter 4, the results of this study indicated that action research interacted with the implementation of new practices in teachers' classroom to predict residual gains in level of concern. However, it was demonstrated that action research was associated with higher residual gains in level of concern when teachers were moderately to heavily involved in implementing new teaching practices. Where teachers were implementing action research but at the same time doing little or nothing new in the classroom, residual gains in level of concern were actually negative.

To a certain degree, these findings are common sense. By definition, action research is rooted in teachers' classrooms; they formulate questions about their own practice and pursue objective answers to these questions. If, however, teachers select and investigate research questions having to do with how they already teach, it is unlikely that they will move forward on the continuum of change. On the other hand, the more teachers are experimenting with and analyzing their use of new strategies, the more likely they are to recognize what works and what doesn't and alter their ways of thinking about learning and teaching. As a result, it is recommended that staff developers focus as much on encouraging teachers to take risks and experiment with teaching as they do on guiding teachers through the research process.

According to the results of this study, participation in solitary, unsupported curriculum development activities are associated with negative residual gains in intensity of concern about change and reported use of teaching practices associated with a change. As a result, it is recommended that teachers not be expected to revise and/or modify curriculum at the same time that they are charged with changing the way they think about and approach teaching. Results of this study suggest that elements of Doyle and Ponder's "practicality ethic" hold true, as certain teachers in this sample found the personal costs (in terms of personal effort and the complexity of the proposed innovation) of revising and developing appropriate standards-based math curriculum to meet the needs of their students too great. These were the same teachers who developed quite negative attitudes about change and actually decreased in the frequency with which they implemented standards-based instruction and assessment.

At the early stages of implementation of an innovation where the goals are to alter teachers' attitudes and practices, therefore, it is recommended that teachers be granted access to curriculum materials that are appropriate for the developmental levels and academic needs of their students. If they find these materials appropriate, teachers will be more likely to actually use them and not become preoccupied with the inappropriateness of the materials. They will also not be charged with the time-consuming, difficult process of designing curriculum that corresponds to a way of teaching and learning with which they are relatively unfamiliar.

The findings of this study indicated that increases in self-efficacy were associated with residual gains in teachers' overall intensity of concern about change. An additional recommendation, therefore, is that efforts be undertaken to increase teachers' self-efficacy, which includes a teachers' general belief in his/her ability to perform a variety of tasks (particularly in the face of adversity), as well as his/her willingness to initiate and persist in the completion of a challenging task. At the present time, this researcher is not aware of any developed self-efficacy training for teachers, although some may no doubt exist. However, it is likely that consistent, one-on-one guidance and support from a coach or consultant around issues with which teachers are grappling will increase their self-efficacy. As a result, it is recommended that administrator identify teachers with low self-efficacy and supply them with someone who can give them feedback about their teaching, help them problem-solve around issues that they are facing, and generally provide them with guidance and moral support around the adoption and implementation of new classroom practices.

An unanticipated finding of this study was that years of teaching experience were significant predictors of residual gains in the frequency with which teachers employed standards-based instructional and assessment practices. This was surprising because much of the previous literature (Berman & McLaughlin, 1977, 1978) has concluded that older teachers are resistant to change and even unable to change. Conventional wisdom supports this notion. Based on the results of this study, however, it is recommended that change initiatives focus more strongly on the older, more experienced teacher. Not only are they highly willing and capable of change, but they may be suitable leaders in change initiatives. At present, many of those chosen to lead their peers in teacher change initiatives are energetic, younger teachers who are assumed to be most knowledgeable about current teaching approaches and strategies. While the younger teachers may or may not be knowledgeable instructional leaders, they are often not well-received by their highly experienced peers. Therefore, it is recommended that more older, experienced teachers be offered the opportunity to lead their colleagues in the journey toward change.

Finally, the results of this study suggested that increased principal support was associated with gains in the frequency with which teachers' used new teaching strategies and practices. Therefore, it is recommended that initiatives that focus on helping teachers change do not ignore the needs of principals who also may need assistance coping with the new demands put on them by a change. Principals themselves may need help learning how to support teachers in ways that will facilitate the adoption of new teaching practices. For this reason, it is suggested that principal support strategies such as those described by Elmore (1996) and implemented in Community District 2, New

York City, be adopted by schools involved in a change process. For example, Elmore explains that, as part of a large scale professional development strategy, District 2 principals regularly visit with other principals in the district and participate in principal support groups where they can share their experience and advice about how to manage change effectively. In addition, principals in District 2 participate in monthly principals' conferences organized around instructional issues. Finally, principals might benefit from coaches or consultants who introduce them to strategies that will make them more supportive of teachers and the change process.

In summary, the implications of the findings of this study include recommendations for change agents and school administrators. It is recommended that those directing a change initiative focus on providing teachers with staff development opportunities that are associated with positive changes in teachers' concern and that lead to increased self-efficacy on the part of teachers. In addition, it is recommended that the staff development planners and change agents recognize that older, more experienced teachers are highly capable of changing the way they think and teach and may even serve as effective leaders to their experienced peers. Finally, it is recommended that the needs of principals not be overlooked, as they, too, may need assistance in providing the support that is associated with changes in teachers' attitudes and instructional practices.

Limitations

The present study has been limited in several ways. First of all, the time period focused on in the study may have been too short for much change to occur in teachers' concerns or practices. As a result, the prediction equations yielded in the analyses and

the conclusions reached may be applicable only to short-term change. The picture for long-term change could conceivably be quite different.

In addition, teacher concerns, teacher and school characteristics, and staff development activities were measured by self-report through the use of surveys and professional development logs. This data collection approach carried with it the risk that respondents were not entirely truthful in their responses. As stated earlier, the reader must be aware that the use of self-report methods does not guarantee truthfulness or objectivity in the information offered by respondents. Nevertheless, this limitation was partly offset by data collected in teacher interviews. If survey and log data had sharply contrasted with interview data, this veracity of the self-report data would have been called into question.

A strong limitation of the present study is its small sample size. Because the study sample (89) represented only 43% of the population (205), it is possible that the teachers in the sample are not representative of the teachers in the population. As a result, it is quite possible that the prediction models yielded in this analysis would not hold true for the population of teachers this sample is meant to represent. In addition, a sample size of 89 is very small for stepwise regression analysis, furthering the danger that the conclusions reached in this study will not generalize beyond this particular sample.

Finally, the use of stepwise regression analysis with several independent variables increases the risk of Type I error. This is particularly true of the analysis involving residual gains in intensity of concern. While adjustments were made to reduce the risk of reaching erroneous conclusions in all analyses, there is no consensus on a single,

preferred adjustment to make or even whether making adjustments completely eliminates this increased risk. If the study sample had been larger and fewer independent variables had been used, this issue would not have been of such great concern.

An additional limitation of this study was the fact that several stepwise regression analyses were conducted on the same sample of study subjects. Consequently, these tests were not independent, thereby increasing the chance of increasing the overall alpha level of these analyses. As stated previously, however, the purpose of this study was exploratory, rather than explanatory, in nature. Consequently, the results of this research should be viewed as tentative.

Suggestions for Future Research

Given the small sample size used in the present study, it is recommended that this study be replicated with a larger sample. This will increase the power of the study and the likelihood that the results obtained are not due purely to sampling error. Replication will also provide valuable data regarding the degree to which the present study is generalizable to all teachers who are involved in a standards-based math reform initiative.

As stated in Chapter 4 and elsewhere, successful implementation of an educational innovation involves changing teachers' attitudes (i.e., their concerns) and their teaching practices. With this in mind, future researchers may want to explore more deeply the nature of the specific variables that were found to predict changes in teachers' concerns and practices:

- Personal variables: self-efficacy and number of years teaching

- Staff development activity variables: implementing standards-based classroom activities, examining student work, action research, and curriculum development
- Organizational variables: principal supportiveness and organizational climate

Additional interviews with teachers will be necessary to explore the nature and importance of these variables more intensively.

It is also recommended that specific models presented in this study be subjected to confirmatory analyses with additional samples. In addition to validating the results of this dissertation study, these procedures would also reduce the complexity of the research by minimizing the number of predictors used in the regression analyses. This will, in turn, increase the power of future multiple regression analyses and eliminate the need to employ stepwise multiple regression techniques, thereby reducing the possibility of an inflated Type I error at the same time. In addition, the use of fewer predictors will likely result in more meaningful and comprehensible results.

Finally, future studies of change in teachers' concerns and practices should investigate other independent variables which may improve the ability to predict teacher change. With the models presented in Chapter 4 only explaining between 13% and 28% of the variance in residual gains in teacher level of concern, intensity of concern, and frequency of use of standards-based instruction and assessment, it is obvious that additional variables are key to predicting and understanding teacher change.

Conclusions

The aim of this dissertation was to explore staff development strategies and contexts that may have a positive impact on teachers' attitudes and practices. The goal of

this study has been met, as it has presented potential models for understanding and predicting teacher change.

Although there are some limitations to the present study, it has opened the door to the use of multiple regression as a means of predicting teacher change from a number of staff development, contextual, and personal variables. In addition, it has yielded information regarding the intensity of staff development activities and the extent to which contextual and personal variables are associated with change.

The results from three major analyses conducted in this study suggested that involvement in staff development (i.e., examining student work, implementing new types of activities, conducting action research, and *not* undertaking single-handed, unsupported attempts at curriculum development), as well as high self-efficacy, are associated with residual gains in level and intensity of concern. To predict residual gains in the use of new teaching strategies, however, contextual factors such as the degree of principal supportiveness (and its relationship with a school or district's organizational climate) entered the picture. In combination with the number of years teaching experience and, again, a lack of involvement in curriculum development, these contextual variables were found to significantly predict residual changes in the frequency of use of new teaching and assessment techniques.

In conclusion, this study has produced evidence suggesting that staff development, contextual, and personal variables predict teacher change, confirming past research that no single variable exists for bringing about changes in teachers' concerns and practices. In fact, teacher change was shown to be associated with eight factors, only

half of which can be considered staff development. Additional factors shown to predict teacher change pertained to teachers as individuals (years experience, self-efficacy) as well as the school as an organization (principal supportiveness/organizational climate). These results represent a gain in the staff development and teacher change literature and, as such, merit further research.

BIBLIOGRAPHY

Abdal-Haqq, I. (2000). Digest Number 89-4, Nature of Professional Development Schools. Washington, DC: ERIC Clearinghouse on Teacher Education. [On-line]. Available: http://www.ericsp.org/digests/digest.html?XP_PUB=digests&XP_RECORD=19990909-173736.

Abdal-Haqq, I. (2000). Digest Number 91-2, Professional Development Schools and Educational Reform: Concepts and Concerns. Washington, DC: ERIC Clearinghouse on Teacher Education. [On-line]. Available: http://www.ericsp.org/digests/digest.html?XP_PUB=digests&XP_RECORD=19990909-173842.

Aschbacher, P. R. (1993). Issues in Innovative Assessment for Classroom Practice: Barriers and Facilitators. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education, University of California, Los Angeles.

Barth, R. S. (1991). Restructuring schools: Some questions for teachers and principals. Phi Delta Kappan, 73 (2), 123-128.

Bennett, B.B. (1987). The Effectiveness Of Staff Development Training Practices: A Meta-Analysis. Doctoral dissertation, University of Oregon.

Bernier, N.R. & McClelland, A.E. (1989). The social context of professional development. In Holly, M.L. & McLoughlin, C.S. (Eds.) Perspectives on teacher professional development. Philadelphia: The Falmer Press.

Berman, P., & McGlaughlin, M. (1977). Federal Programs Supporting Educational Change, Vol. VII: Factors Affecting Implementation and Continuation. Santa Monica, CA: The Rand Corporation.

Berman, P. & McGlaughlin, M. (1978), Federal Programs Supporting Educational Change, Vol. VIII: Implementing and Sustaining Innovations. Santa Monica, CA: The Rand Corporation.

Blackman, C.A. (1989). Issues in professional development: the continuing agenda. In Holly, M.L. & McLoughlin, C.S. (Eds.) Perspectives on teacher professional development. Philadelphia: The Falmer Press.

Borko, H. (1997). New forms of classroom assessment: Implications for staff development. Theory into Practice, 36 (4), 231-238).

Borko, H., & Putnam, R. (1996). Learning to teach. In R. Calfee & D. Berliner (Eds.), Handbook of educational psychology (pp.673-708). New York: Macmillan.

Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. Teaching & Teacher Education, 13 (3), 259-278.

Bredeson, P. V. & Scribner, J. P. (2000). A statewide professional development conference: Useful strategy for learning or inefficient use of resources? Educational Policy Analysis Archives, 8 (13). [On-line]. Available: <http://epaa.asu.edu/epaa/v8n13.html>.

Burnes, A. (November 2-5, 1995). Teachers' Voices: Curriculum Design and Change. Paper presented at the Annual Meeting of the Japan Association of Language Teachers, Nagoy, Japan.

Busher, L. A. (1994). The Effects of Peer Coaching on Elementary School Teachers. Paper presented at the Annual Meeting of the Eastern Educational Research Association, Sarasota, FL.

Carnegie Forum on Education and the Economy, Task Force on Teaching as a Profession. (1986). A Nation prepared: Teachers for the 21st Century. New York: Author.

Clinard, L.M., Ariav, T., Beeson, R., Minor, L., Dwyer, M. (1995). Cooperating Teachers Reflect upon the Impact of Coaching on Their Own Teaching and Professional Life. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco. Available online: <http://www.gse.uci.edu/doehome/DeptInfo/Faculty/Clinard/AERA.HTML>.

Clune, W. H. (1991). Systemic Educational Policy. Madison, WI: Wisconsin Center for Educational Policy, University of Wisconsin-Madison.

Cohen, J. (1992). A power primer. Psychological Bulletin, 112 (1), 155-159.

Crandall, D. P. & Loucks, S. F. (1983). A Roadmap for School Improvement. Executive Summary of the Study of Dissemination Efforts Supporting School Improvement. People, Policies, and Practices: Examining the Chain of School Improvement, Volume X. Washington, DC: Department of Education.

Daresh, J.C. (1985). Research Trends in Staff Development and Inservice Education. Paper presented at the Annual Meeting of the Mid-Western Educational Research Association (Chicago, IL, October 1985). ED292180.

Darling-Hammond, L. (1995). Policies that support professional development in an era of reform. Phi Delta Kappan, 76 (8), 597-604.

Darling-Hammond, L. (1995). Rethinking teacher leadership through professional development schools. Elementary School Journal, 96 (1), 87-106.

Darling-Hammond, L. (1996). The quiet revolution: Rethinking teacher development. Educational Leadership, 53 (6), 4-10.

Darling-Hammond, L. (1998). Teacher learning that supports student learning. Educational Leadership, 55 (5), 6-11.

Durzo, J., Florini, B. M. (1984). Evaluating instructional development projects: Guidance from the literature on change. Instructional Development: The State of the Art, Vol. II.

Edelfelt, R. & Lawrence, G. (1975). In-service education: The state of the art. In Edelfelt, R. A. & Johnson, M., Eds. (1975). Rethinking In-Service Education. Washington, DC: National Education Association.

Elmore, R.F. (1995). Structural reform and educational practice. Educational Researcher, 24 (9), 23-26.

Elmore, R.F. (1996). Getting to scale with good educational practice. Harvard Educational Review, 66 (1), 1-26.

Elmore, R.F. and Consortium for Policy Research in Education (1996). Staff Development and Instructional Improvement: Community District 2, New York City. Paper prepared for the National Commission on Teaching and America's Future.

Elmore, R.F. & Burney, D. (1997). School Variation and Systemic Instructional Improvement in Community School District #2, New York City. Paper prepared under the sponsorship of the High Performance Learning Communities Project at the Learning Research and Development Center, University of Pittsburgh.

Fenstermacher, G. & Berliner, D. (1985). Determining the value of staff development. The Elementary School Journal, 85 (3), 282-314.

Fielding, G. D. & Schalock, H. D. (1985). Promoting the Professional Development of Teachers and Administrators. ERIC/CEM School of Management

Digest Series, Number 31. Eugene, OR: ERIC Clearinghouse on Educational Management, Oregon University.

Fitch, M.E. & Kopp, O.W. (1990). Staff Development for the Practitioner. Springfield, IL: Charles Thomas Publisher.

Flexer, R. J., Cumbo, K., Borko, H., Mayfield, V., & Marion, S. (1995). How “messing about” with performance assessment in mathematics affects what happens in classrooms. CSE Technical Report 396. Los Angeles: National Center for Research in Evaluation, Standards, and Student Testing (CRESST).

Fullan, M. G. (1982). Implementing Educational Change: Progress at Last. Paper presented at the National Invitational Conference, “Research on Teaching: Implications for Practice,” Warrenton, VA.

Fullan, M. G. (1999). Change Forces: The Sequel. Bristol, PA: Falmer Press.

Fullan, M. G. (1993). Change Forces: Probing the Depths of Educational Reform. London: Falmer Press.

Fullan, M. G. (1994). Change Forces: Probing the Depths of Educational Reform. School Development and the Management of Change Series: 10. Bristol, PA: Falmer Press.

Fullan, M.G. (1995). The limits and potential of professional development. In T. R. Guskey & M. Huberman (Eds.), Professional Development Education: New Paradigms and Practices (pp.253-267). New York: Teachers College Press.

Gage, N. L. (1984). What do we know about teaching effectiveness? Phi Delta Kappan, 66 (2), 87-83.

Gage, N. L. & Needels, M. C. (1989). Process-product research on teaching: A review of criticisms. The Elementary School Journal, 89 (3), 253-300.

Gall, M.D. & Vojtek, R.O. (1994). Planning for Effective Staff Development: Six Research-Based Models. Eugene, OR: ERIC Clearinghouse on Educational Management. ED372464.

Gliessman, D. H., Pugh, R. C., Dowden, D. E. & Hutchins, T. F. (1988). Variables influencing the acquisition of a generic teaching skill. Review of Educational Research, 58 (1), 25-46.

Goldenberg, C. & Gallimore, R. (1991). Changing teaching takes more than a one-shot workshop. Educational leadership, 49 (3), 69-72.

Good, T. L. & Grouws, D. A. Increasing teachers' understanding of mathematical ideas through inservice training. Phi Delta Kappan, 68 (10), 778-783.

Gordon, L. (1974). Patterns of Effective Inservice Education. A State of the Art Summary of Research on Materials and Procedures for Changing Teacher Behaviors in Inservice Education. Tallahassee, FL: Florida State Dept. of Education.

Griffin, G. (1989). The school in society and the societal organization of school: Implications for staff development. In Holly, M.L. & McLoughlin, C.S. (Eds.) Perspectives on teacher professional development. Philadelphia: The Falmer Press.

Guskey, T. R. & Huberman, M. (1995). Introduction. In T. R. Guskey & M. Huberman (Eds.), Professional Development Education: New Paradigms and Practices (pp. 1-6). New York: Teachers College Press.

Guskey, T. R. (1997). Research needs to link professional development and student learning. Journal of Staff Development, 18 (20). Available: <http://www.nsd.org/library/jsd/jsdgusk.html>

Guskey, T. R. (1995). Professional development in education: In search of the optimal mix. In T. R. Guskey & M. Huberman (Eds.), Professional Development Education: New Paradigms and Practices (pp. 114-131). New York: Teachers College Press.

Guskey, T. R. (1986). Staff development and the process of teacher change. Educational Researcher, 15 (5), 5-12.

Guskey, T. R. (1994). Results-oriented professional development: In search of an optimal mix of effective practices. Journal of Staff Development, 15 (4), 42-50.

Hall, G. E. & Hord, S. M. (1987). Change in Schools: Facilitating the Process. Albany, NY: State University of New York Press.

Hall, G. E., George, A., & Rutherford, W.L. (1977). Measuring Stages of Concern about the Innovation: A Manual for the Use of the SoC Questionnaire. Austin, TX: Research and Development Center for Teacher Education.

Halpin, A.W. (1963). The Organizational Climate of Schools. Washington, DC: U.S. Office of Education, Department of Health, Education, & Welfare.

Havelock, R. G. (1973). The Change Agent's Guide to Innovation in Education. Englewood Cliffs, NJ: Educational Technology Publications.

Higuchi, C. (1993). Performance -based assessments and what teachers need. CSE Technical Report 362. Los Angeles: National Center for Research in Evaluation, Standards, and Student Testing (CRESST).

Holly, M.L.H. (1989). Teacher professional development: Perceptions and practices in the USA and England. In Holly, M.L. & McLoughlin, C.S. (Eds.) Perspectives on teacher professional development. Philadelphia: The Falmer Press.

Holmes Group. (1986). Tomorrow's Teachers. East Lansing, MI: Author.

Horst, P. (1941). The role of prediction variables which are independent of the criterion. In P. Horst (Ed.), The prediction of personal adjustment. Social Research Bulletin, 48, 431-436.

Horst, P. (1966). Psychological Measurement and Prediction. Belmont, CA: Wadsworth.

Hyde, A.A. & Pink, W.T. (1992). Thinking about effective staff development. In Pink, W.T. & Hyde, A.A. (Eds.). Effective Staff Development for School Change. Norwood, NJ: Ablex Publishing Corporation.

Jaccard, J., Turrisi, R., Wan, C.K. (1990). Interaction Effects in Multiple Regression. Newbury Park, CA: Sage Publications.

Joyce, B. & Showers, B. (1980). Improving inservice training: The messages of research. Educational Leadership, 37 (5), 379-385.

Joyce, B. & Showers, B. (1982). The coaching of teaching. Educational Leadership, 40 (1), 4-10.

Joyce, B. & Showers, B. (1983). Power in Staff Development through Research on Training. Alexandria, VA: Association for Supervision and Curriculum Development.

Joyce, B. & Showers, B. (1995). Student Achievement through Staff Development: Fundamentals of School Renewal, 2nd Ed. White Plains, NY: Longman.

Joyce, B. R., McNair, K. M., Diaz, R., & McKibbin, M. D. (1976). Interviews: Perceptions of professionals and Policy Makers. Stanford, CA: Stanford Center for Research and Development in Teaching, Stanford University.

Kerman, S. (1979). Teacher expectations and student achievement. Phi Delta Kappan, 60 (10), 716-718.

Killion, J. & Harrison, J. (1997). The multiple roles of staff developers. Journal of Staff Development, 18 (3). Available online: <http://www.nsdc.org/library/jsd/killion183.html>.

Kimpston, R. D. & Rogers, K. B. (1987). Collaborative Curriculum Planning: Does the Process Contribute to Changes in Attitudes toward Curriculum? Paper presented at the Annual Meeting of the American Educational Research Association, Washington, DC.

Kirkpatrick, D.L. (1959). Techniques for evaluating training programs. Journal of the American Society for Training & Development, 13 (11), 3-9.

Kirkpatrick, D.L. (1977). Evaluating training programs: Evidence vs. proof. Training & Development Journal, 31 (10), 9-12.

Kubler-Ross, E. (1969). On Death and Dying. New York: Macmillan Publishing Co.

Lambert, L. (1988). Staff development redesigned. Educational Leadership, 45 (8), 665-668.

Lawrence, G. (1974). Patterns of Effective Inservice Education. A State of the Art Summary of Research on Materials and Procedures for Changing Teacher Behaviors in Inservice Education. Tallahassee: Florida Department of Education.

Lewin, K. (1958). Group decision and social change. In E. E. Maccoby, T. M. Newcomb, & E. L. Hartley (Eds.) Readings in Social Psychology (3rd ed.). New York: Henry Hold and Company.

Licht, M.H. (1997). Multiple regression and correlation. In Grimm, L.G. & Yarnold, P.R., Eds. Reading and Understanding Multivariate Statistics. Washington, DC: American Psychological Association.

Lieberman (1995). The Work of Restructuring Schools: Building from the Ground Up. New York: Teachers College Press.

Lippit, R., Watson, J., & Westley, B. (1958). The Dynamics of Planned Change. New York: Harcourt, Brace and World.

Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of School Success. American Educational Research Journal, 19 (3), 325-340.

Little, J.W. et al. (1987). Staff Development In California. Public And Personal Investments, Program Patterns, And Policy Choices. California Postsecondary Education

Commission Report. CA: Far West Lab For Educational Research And Development & Policy Analysis For California Education. ED300342.

Little, J.W. (1993). Teachers' professional development in a climate of educational reform. Educational Evaluation and Policy Analysis, 15 (2), 129-151.

Mann, D. (1986). Authority and school improvement: An essay on "Little King" leadership. Teachers College Record, 88 (1), 41-51.

Manning, B. A. (1976). The "Trouble Shooting" Checklist for School-Based Settings (Manual). Austin, TX: Research and Development Center for Teacher Education, Texas University.

Massarella, J.A. (1980). Synthesis of research on staff development. Educational Leadership, 38 (2), 182-185.

McGlaughlin, M. & Berman, B. (1977). Retooling staff development in a period of retrenchment. Educational Leadership, 35 (3), 191-194.

McGlaughlin, M. W. & Marsh, D. D. (1978). Staff Development and School Change. Teachers College Record, 80 (1), 69-94.

Miles, M. B., Farrar, E. & Neufeld, B. (1983). The Extent of Adoption of Effective Schools Programs. Vol. 11. Review of Effective Schools Programs. Washington, DC: Department of Education.

Mohlman, G. G., Coladarci, T., & Gage, N. L. (1982). Comprehension and attitude as predictors of implementation of teacher training. Journal of Teacher Education, 23 (1), 31-36.

Montgomery, B. and Way, W. L. (December 1995). Lost and Found Voices in the Process of Curriculum Change. Paper presented at the Annual Meeting of the American Educational Research Association, Denver, CO.

Murphy, C. U. (May 1991). Lessons from a Journey into Change. Educational Leadership, 63-67.

Murphy, C. U. & Lick, D.W. (1998). Whole-Faculty Study Groups. Thousand Oaks, CA: Corwin Press, Inc.

Neuman, G. A., Edwards, J. E., & Raju, N. S. (1989). Organizational development interventions: A meta-analysis of their effects on satisfaction and other attitudes. Personnel Psychology, 42 (3), 461-483.

Novick, R. (1996). Actual schools, possible practices: New directions in professional development. Education Policy Analysis Archives, 4 (14). Available: <http://olam.ed.asu.edu/epaa/v4n2.html>.

Orlich, D. C. (1988). An Emergent Paradigm for Staff Development. Paper presented at the Annual Conference of the National Council of States on Inservice Education, New Orleans, LA.

Osborne, B. (1993). Understanding Change in a Time of Change. Paper presented at the Annual Conference on Creating the Quality School, Oklahoma City, OK.

Peca, K. (1994). Focusing on the Individual Change Process in School Restructuring. Keynote address at the Texas Education Agency Institute on Reaching All Students: Building School District Capacity, Austin, TX.

Pedhazur, E.J. (1997). Multiple Regression in Behavioral Research: Explanation and Prediction, 3rd Ed. NY: Harcourt Brace College Publishers.

Pink, W. T. (1986). Facilitating change at the school level: A missing factor in school reform. The Urban Review, 19-30.

Ponder, G. & Doyle, W. (1977). Teacher Practicality and Curriculum Change: An Ecological Analysis. Paper presented at the Annual Meeting of the American Educational Research Association, New York, NY.

Professional Development Schools at a Glance. (2000). Washington, DC: American Association of Colleges for Teacher Education. [On-line]. Available: http://www.aacte.org/pds/html?XP_PUB=pds&XP_RECORD=20000214-155852.

Professional Development Schools: Frequently Asked Questions. (2000). Washington, DC: American Association of Colleges for Teacher Education. [On-line]. Available: http://www.aacte.org/faq/html?XP_PUB=faq&XP_RECORD=20000424-161549.

Rieley, J. B. (1997). Environments for change. ERIC Opinion Papers. (ED 411 007).

Robbins, P. & wolfe, P. (1987). Reflections on a Hunter-based staff development project. Educational Leadership, 44 (5), 56-61.

Rogers, E. M. & Rogers, R. A. (1976). Communication in Organizations. New York: Free Press, 1976.

Rogers, E. M. & Shoemaker, F. F. (1971). Communication of Innovations: A Cross-Cultural Approach. New York: Free Press, 1971.

Sagor, R. (March 1991). What Project LEARN reveals about collaborative action research. Educational Leadership, 6-10.

Sagor, R. D. & Curley, J. L. (1991). Collaborative Action Research: Can It Improve School Effectiveness? Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Saljo, R. (1997). Self-report in educational research. In Keeves, J.P. (Ed.). Educational Research, Methodology, and Measurement: An International Handbook, 2nd Ed. (pp. 101-106). NY: Elsevier Science Ltd.

Saxe, G. B., Franke, M. L., Gearhart, M., Howard, S., & Crockett, M. (1997). Teachers' shifting assessment practices in the context of educational reform in mathematics. CSE Technical Report 471. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Shepard, L. A. (1995). Using assessment to improve learning. Educational Leadership 52 (3), 38-45.

Shepard, L. A. (1997). Insights Gained from a Classroom-Based Assessment Project. CSE Technical Report 451. National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Graduate School of Education & Information Studies, University of California, Los Angeles.

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The Self-Efficacy Scale: Construction and Validation. Psychological Reports, 51, 663-671.

Showers, B., Joyce, B., & Bennett, B. (1987). Synthesis of research on staff development: A framework for future study and a state-of-the-art analysis. Educational Leadership, 45 (3), 77-87.

Sparks, D. (October 1996). Standards-based reform requires standards-based staff development. The Developer. [On-line]. Available: <http://www.nsd.org/library/developer/dev10-96sparks2.html>.

Sparks, D. & Loucks-Horsley, S. (1990). Models of staff development, in Houston, R.W. (Ed.). Handbook of Research on Teacher Education. NY: Macmillan Publishing Co.

Sparks, G.M. (1983). Synthesis of research on staff development for effective teaching. Educational Leadership, 41 (3), 65-72.

Sparks, G. M. (1986). The effectiveness of alternative training activities in changing teacher practices. American Educational Research Journal, 23 (2), 217-225.

Sparks, G. M. & Bruder, S. (1987). Before and after peer coaching. Educational Leadership, 45 (3), 54-57.

Sparks, G., Nowakowski, M., Hall, B., Alec, R., & Imrick, J. (1985). School improvement through staff development. Educational Leadership, 42 (6), 59-61.

Staff Development. Research Action Brief Number 10. (1980). Eugene, OR: ERIC Clearinghouse on Educational Management. ED 189679.

Stallings, J.A. & Kowalski, T. (1990). Research on Professional Development Schools, in Houston, R.W. (Ed.). Handbook of Research on Teacher Education. NY: Macmillan Publishing Co.

Stallings, J. & Mohlman, G. (1981). School policy, leadership style, teacher change, and student behavior in eight schools. Final Report, National Institute of Education, Washington, DC.

Stout, R. T. (1996). Staff development policy: Fuzzy choices in an imperfect market. Education Policy Analysis Archives, 4 (2). Available: <http://olam.ed.asu.edu/epaa/v4n2.html>.

Swenson, T. L. (1981). The state of the art in inservice education and staff development. Journal of Research and Development in Education, 15 (1), 2-7

Tabachnick, B.G. & Fidell, L.S. (1996). Using Multivariate Statistics, 3rd Ed. NY: HarperCollins College Publishers.

Teitel, L. & Abdal-Haqq, I. (2000). Assessing the Impacts of Professional Development Schools. Washington, DC: American Association of Colleges for Teacher Education.

Thompson, B. (1995). Stepwise regression and stepwise discriminant analysis need not apply here: A guidelines editorial. Educational and Psychological Measurement, 55 (4), 525-534.

Tutt, B. & Carter, S. (1999). Understanding and Using Change Forces. Paper presented at the Annual Meeting of the American Association of Colleges for Teacher Education, Washington, DC.

Tye, K. A. & Tye, B. B. (1984). Teacher isolation and school reform. Phi Delta Kappan, 65 (5), 319-322.

United States: The quest for standards, accountability, and excellence. (1998). In Staying Ahead: In-Service Training and Teacher Professional Development. Paris: OECD.

Wade, R. (1985). What makes a difference in inservice teacher education? A meta-analysis of research. Educational Leadership, 42 (4), 48-54.

Ward, B. & Tikunoff, W. (1981). The relationship between inservice training, organizational structure, and school climate. Inservice, 7-8.

Watts, H. (1985). When teachers are researchers, teaching improves. Journal of Staff Development, 6 (2), 118-127.

Waugh, R. F. & Punch, K. F. (1987). Teacher receptivity to systemwide change in the implementation stage. Review of Educational Research, 57 (3), 237-254.

Weatherly, R. & Lipsky, M. (1977). Street-level bureaucrats and institutional innovation: Implementing special education reform. Harvard Educational Review, 47 (2), 171-197.

Wertheimer, R. & Zinga, M. (1997). Attending to the Noise: Applying Chaos Theory to School Reform. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.

Wideen, M.F. (1987). Perspectives on staff development. In Wideen, M.F. & Andrews, I. (Eds.). Staff Development for School Improvement. Philadelphia: The Falmer Press.

Wilkinson, L. & Dallal, G.E. (1981). Tests of significance in forward selection regression. Technometrics, 23 (4), 377-380.

Wise, A. E. (1991). On teacher accountability. In Voices from the Field (pp. 23-24). Washington, DC: William T. Grant Foundation Commission for Work, Family and Citizenship and Institute for Education Leadership.

Wood, F.H. & Thompson, S.R. (1993). Assumptions about staff development based on research and best practice. Journal of Staff Development, 14 (4), 52-57.

Zaltman, G., Duncan, R. & Holbeck, J. (1973). Innovations and Organizations.
New York: John Wiley and Sons.

APPENDIX 1: Teacher Pre- and Post-Surveys

APPENDIX 2: Professional Development Activity Log

APPENDIX 3: Student Survey

APPENDIX 4: Interview Protocol

